

Political Economy of State Capacity



Mats-Philip Ahrenshop
Trinity College
University of Oxford

Thesis submitted in partial fulfillment of the requirements for the
degree of

Doctor of Philosophy

Michaelmas Term 2023

To my parents Frauke and Dietrich, for they have opened up the world
to me

Acknowledgements

This thesis is the product of the help, encouragement and support from many people. First and foremost, I wish to thank my supervisors David Rueda and Spyros Kosmidis for their never-ending patience, brilliance and insights. I could not have asked for better mentors.

I benefited greatly from funding of the Economic and Social Research Council, Trinity College Oxford and the International Centre for Tax and Development at the University of Sussex.

Various parts of the thesis benefited from the comments, suggestions and criticisms of Noah Bacine, Sarah Brierley, Leonardo Carella, Amrita Dhillon, David Doyle, Raymond Duch, Ashrakat Elshehawy, Stephen Fisher, Tanushree Goyal, Robin Harding, Adrienne LeBas, César Fuster Llamazares, Tanvi Ravel Mehta, Francesco Raffaelli, Ryan Rholes, James Robinson, Vanessa van den Boogaard, Jonathan Weigel as well as participants at the 2021 EUI Graduate Students Workshop, the Nuffield CESS Colloquium, the DPIR Research in Progress Seminar and the IAST-OxPo Political Economy Conference. I am also indebted to my amazing co-authors Augustin Bergeron, Laura Paler, Gabriel Tourek and Jonathan Weigel; I cannot say enough how much I have learned from you.

I am grateful to many people who helped with data collection, including the amazing team at ODEKA in Kananga, DRC, as well as the wonderful teams at Bilendi & Respondi and the Nuffield Centre for Experimental Social Sciences.

Thank you to my friends Ash, César, Christine, Francesco, Leo, Marta, Raquel and Scott for the long coffee breaks in the department and the constant encouragement that life goes on. Thank you to Henry and Kaya, for you have made all the difference to my mental well-being. Finally, I wish to thank Eddy for everything that she has done for me, big and small, every single day in our beautiful relationship.

Abstract

How does society react politically when the state tries to increase its capacity? This thesis presents three experimental studies that shed light on this question by examining the micro-behavioral consequences of such increases in state capacity. The first article explores how delegating tax collection to local elites such as city chiefs in the Democratic Republic of the Congo affects citizen demand for accountability, both from the state and the chiefs. It examines data collected from a randomized controlled trial in the city of Kananga where neighborhoods were assigned to either state or chief tax collection, and citizens had a collective action opportunity to hold both accountable through the request of citizen meetings. The results show that delegating tax collection to local elites does not impede citizens' ability or willingness to hold both chiefs and the state accountable for policy outcomes. The second article asks whether there are electoral costs associated when candidates for political office promise to fight corruption instead of improving economic growth. The study draws on evidence from two conjoint experiments conducted in India and the United Kingdom where respondents could choose between different hypothetical candidates who varied along multiple dimensions. The results suggest that promising to fight corruption as a central policy platform may be electorally punished by voters compared to promising to improve the economy. The third article investigates the bottom-up accountability pressures in response to earmarking tax revenue for particular public expenditures. It does so by leveraging variation in the extent of citizens' beliefs about the prevalence of earmarking in an online survey experiment conducted in Ghana. Results show that earmarking revenue does not have positive accountability effects by virtue of increasing citizens' incentives to engage in politics. Overall, this thesis contributes original data, novel evidence and new theoretical frameworks to the study of state-society relations and the political economy of state capacity.

Contents

1	Introduction	1
2	Delegating Tax Collection Does Not Adversely Affect Demand for Accountability: Evidence from an Experiment in the D.R. Congo	9
2.1	Introduction	9
2.2	Theoretical Framework	13
2.2.1	Voice in Response to Tax Collection Delegation	13
2.2.2	Scope Conditions	16
2.3	Experimental Design and Implementation	18
2.3.1	Setting	18
2.3.2	Treatments	18
2.3.3	Outcome Measures	19
2.3.4	Analysis	22
2.4	Results	23
2.4.1	First-Stage Results	23
2.4.2	Main Results	24
2.4.3	Alternative Explanations	30
2.5	Discussion	34
3	Do Anti-Corruption Policies Attract Votes? Experimental Evidence from India and the U.K.	37
3.1	Introduction	37
3.2	Theoretical Framework	41
3.2.1	The Issue Salience of Corruption in Election Campaigns	41
3.2.2	Expected Benefits from Fighting Corruption	42
3.3	Experimental Design and Implementation	45
3.3.1	Setting	45
3.3.2	Treatments	46
3.3.3	Outcome Measures	49

3.3.4	Analysis	50
3.4	Results	52
3.4.1	Anti-Corruption Campaign Platforms	52
3.4.2	Anti-Corruption Platforms, Costs and Effectiveness	54
3.4.3	Budgetary Salience of Anti-Corruption Policies	56
3.4.4	Treatment Effect Heterogeneity	57
3.5	Discussion	62
4	Tax Earmarking and Political Participation: Theory and Evidence from Ghana	65
4.1	Introduction	65
4.2	Background and Motivating Example	69
4.3	Theoretical Framework	70
4.3.1	Preferences	70
4.3.2	Incentives	71
4.3.3	Main Assumptions	78
4.3.4	Alternative Mechanisms	80
4.4	Experimental Design and Implementation	82
4.4.1	Setting	82
4.4.2	Treatments	83
4.4.3	Outcome Measures	84
4.4.4	Analysis	85
4.5	Results	86
4.5.1	Main Results	86
4.5.2	Alternative Mechanisms	89
4.5.3	Explaining Null Results	91
4.5.4	Heterogeneity	93
4.6	Discussion	93
5	Conclusion	96
	Appendices	99
A	Delegating Tax Collection Does Not Adversely Affect Demand for Accountability: Evidence from an Experiment in the D.R. Congo	100
A.1	Additional Information on Experimental Design	100
A.1.1	Intervention and Data Collection Activities	100

A.1.2	Audit Request Form	101
A.1.3	Design Diagnostics: Balance and Interference	103
A.2	Appendices for Main Results	107
A.2.1	Robustness Checks	107
A.2.2	Additional Analyses	109
A.2.3	Mechanisms	110
A.3	Appendices for Alternative Explanations	113
A.3.1	Meetings and Genuine Accountability	113
A.3.2	Tax Collection Salience	119
A.3.3	Chief Tax Collection Targeting	121
A.4	Exploratory Heterogeneous Effects	123
A.5	Discussion of Differences with Pre-Analysis Plan	124
B	Do Anti-Corruption Policies Attract Votes? Experimental Evidence from India and the U.K.	128
B.1	Additional Information on Experimental Design	128
B.1.1	UK Survey Instrument	128
B.1.2	India Survey Instrument	132
B.1.3	UK Subject Sample Characteristics	136
B.1.4	India Subject Sample Characteristics	136
B.1.5	Design Diagnostics: Profile Order and Carryover Effects	137
B.2	Appendices for Main Results	140
B.3	Discussion of Differences with Pre-Analysis Plan	149
C	Tax Earmarking and Political Participation: Theory and Evidence from Ghana	150
C.1	Additional Information on Experimental Design	150
C.1.1	Survey Instrument	150
C.1.2	Design Diagnostics	155
C.2	Appendices for Main Results	156
C.3	Discussion of Differences with Pre-Analysis Plan	159
	Bibliography	160

List of Figures

2.1	Salience of Tax Collection by City Chiefs	24
2.2	Effect of Chief Collection on Tax Compliance and Revenue	25
2.3	Relationship Between Tax Revenue and Form Submission in State and Chief Neighborhoods	28
3.1	AMCEs of the Candidate Choice Experiment	54
3.2	ACIEs for Varying Policy Costs	55
3.3	ACIEs for Varying Policy Effectiveness	57
3.4	Percentage Allocated for a Hypothetical Government Budget	58
3.5	AMCEs Conditional on Household Income	60
3.6	AMCEs Conditional on Incumbent Party Support	61
A.1	Audit Request Form	102
A.2	Recall of Tax Payments Remitted to Which Actor	110
A.3	Number of Demands on Local Actors	111
A.4	Responsibility for Local Public Goods Provision	112
A.5	Meeting Request Forms Comments	116
A.6	Most Common Topics in Citizen Meetings	117
A.7	Most Common Words in Citizen Meetings	118
B.1	Test for Profile Order Effects	137
B.2	Test for Carryover Effects	138
B.3	Display Frequencies of Conjoint Levels	139
B.4	Difference in AMCEs by Income	141
B.5	Difference in AMCEs by Incumbent Support	142
B.6	Marginal Means	143
B.7	Marginal Means for Issue/Costs Interaction	144
B.8	Marginal Means for Issue/Effectiveness Interaction	145
B.9	Subgroup Marginal Means by Income	146
B.10	Subgroup Marginal Means by Incumbent Support	147

B.11 AMCEs by Household Income: Binary	148
C.1 25 Most Written Words in Open-Ended Question	158

List of Tables

2.1 Effects of Chief Tax Collection on Demand for Citizen Meetings . . .	26
2.2 Effects of Chief Tax Collection on Expectations about Chiefs and Gov- ernment	29
2.3 Effects of Tax Campaign Timing on Demand for Citizen Meetings . .	31
2.4 Effects of Chief Tax Collection on Fear of Coercion and Chief Cooptation	32
2.5 Effects of Chief Tax Collection: Payment Propensity	34
3.1 Exemplary Conjoint Table	47
3.2 India Conjoint: Attribute Levels	49
4.1 Hypotheses Tests and Predictions of Treatment Effects	87
4.2 Effect of Tax Earmarking on Political Participation	88
4.3 Effect of Tax Earmarking on Political Participation: Mechanisms . .	89
4.4 Effect of Tax Earmarking on Political Participation: Alternative Mech- anisms	90
4.5 Effect of Tax Earmarking on Political Participation: Explanations for Null Results	92
4.6 Effect of Tax Earmarking on Political Participation: Heterogeneity . .	94
A.1 Treatment Allocation in the 2018 Property Tax Campaign	101
A.2 Timing of all Activities and Data Collection	101
A.3 Balance Tests for Covariates	104
A.4 Balance Tests For Covariates Among Flier Recipients	105
A.5 Extent of Interference in Treatments	106
A.6 Effects of Chief Tax Collection on Demand for Citizen Meetings: More Controls	107
A.7 Effects of Chief Tax Collection on Demand for Citizen Meetings: Al- ternative Strata	108
A.8 Effect of Chief Tax Collection on Civic Participation	109

A.9	LDA Topic Model for Comments in Citizen Meetings	114
A.10	Meeting Attendance Rates by Form Submission	115
A.11	Effects of Chief Tax Collection: Tax Prime	119
A.12	Absolute and Relative (Row) Frequencies of Form Submissions by Treatment Groups	119
A.13	Effects of Chief Tax Collection: Amount of Taxes Collected	120
A.14	Effects of Chief Tax Collection on Ethnic Salience	121
A.15	Effects of Chief Tax Collection: Main Tribe	122
A.16	Effects of Chief Tax Collection: Respondent Gender	123
A.17	Effects of Chief Tax Collection on Trust in Institutions	124
A.18	Effects of Chief Tax Collection on Additional Outcomes in Endline I .	125
A.19	Effects of Chief Tax Collection on Additional Outcomes in Endline II	126
A.20	Effects of Chief and Govt. Collection on Demand for Citizen Meetings: Comparison to Pure Control Group	127
B.1	UK Conjoint: Attribute Levels	130
B.2	India Conjoint: Attribute Levels	133
B.3	UK Subject Sample Characteristics	136
B.4	India Subject Sample Characteristics	136
B.5	India AMCEs	140
B.6	UK AMCEs	140
C.1	Balance Tests for Covariates	155
C.2	Factual Manipulation Checks	156
C.3	Effect of Tax Earmarking on Political Participation: Covariate Adjust- ment	156
C.4	LDA Topic Model for Comments in Open-Ended Question	157

Chapter 1

Introduction

The ability of the state to enforce laws, collect taxes, and protect property rights is central to the study of political science. It is at the heart of much scholarship today that seeks to explain the causes and consequences of state capacity in many fields of comparative politics ranging from the study of conflict, violence and war (Centeno, 1997; Acemoglu, Robinson and Santos, 2013; Sanchez de la Sierra, 2020) as well as economic development (Besley and Persson, 2010; Acemoglu, García-Jimeno and Robinson, 2015; Bardhan, 2016) to the emergence of democracy (North and Weingast, 1989; Olson, 1993; Scheve and Stasavage, 2012). Throughout history, states have tried to build capacity in response to both external threats and internal pressures. They did so by setting up large, modern bureaucracies that were able to deliver impartial law enforcement and equitable and efficient collection of taxes (Besley and Persson, 2009; Fukuyama, 2011; Pomeranz and Vila-Belda, 2019).

What is often neglected is society's political and behavioral response to such increases in state capacity. How do citizens react when the state tries to increase its power? In modern democracies today, citizens are expected to hold government accountable for attempts at increasing its capacity. Canonical models of electoral accountability emphasize the direct accountability relationships between citizens as voters and elected officials in and around democratic elections (Key, 1966; Fiorina, 1981; Fearon, 1999). Formal accountability through (electoral) reward and punishment is then supposed to lead to better governance (Ashworth, 2012; Christensen, Hartman and Samii, 2021*a*).

Increasingly, scholars are asking about this balance of power between state and society. Migdal (1988) shows that state capacity depends on the strength of civil society in a zero-sum fashion. By contrast, Acemoglu and Robinson (2019) argue that it is only when civil society is strong enough to reign in the state's attempt at building capacity that effective state capacity emerges. Stasavage (2020) explores the

possibility that where a strong state bureaucracy developed first, autocratic forms of governance would emerge instead of democratic ones.

However, we lack granular as well as causally identified evidence on the relationship between citizens and the state. In particular, we know little about how to characterize citizens' political and behavioral reactions to different types of building state capacity at the micro level. Two main areas of research have received recent scholarly attention and have produced two main findings: that taxation, which represents a rise in state capacity, engenders citizen bottom-up pressure for accountability (Ross, 2001; Paler, 2013; Weigel, 2020), and that political corruption, which represents a decline in state capacity, oftentimes does not face electoral accountability pressures by voters (Muñoz, Anduiza and Gallego, 2016; De Vries and Solaz, 2017; Klačnja, Lupu and Tucker, 2020).

In this thesis, I am asking new questions. How exactly does taxation, both in terms of its tax-benefit linkage as well as the act of collecting it, catalyze political participation and accountability? Is fighting corruption electorally rewarded by voters? How does increasing state capacity—by taxing its population in new ways or by fighting corruption—affect civil society?

These questions are especially pertinent in developing countries and thus in contexts of weak or emerging state capacity. These contexts are often characterized by low levels of tax revenue, weak property rights protection, high levels of informality, natural resource wealth as well as dysfunctional political institutions (Besley and Persson, 2014; Blattman, Hartman and Blair, 2014; Hartman, Blair and Blattman, 2021). This could have implications for the dynamics of the relationship between state and society, especially in the domains of tax collection and fighting corruption (Weigel, 2020; Weigel and Ngindu, 2021; Balán et al., 2022).

This thesis intends to shed light on these questions with experimental evidence from a range of developing countries—the Democratic Republic of the Congo, Ghana, and India—on the effects of state capacity building on citizens' political behavior. Chapter 2 asks whether delegating tax collection to local elites such as city chiefs in the DRC has adverse consequences for bottom-up accountability pressure by citizens. Chapter 3 investigates in India and the U.K. whether promising to fight corruption as a political candidate generates electoral rewards by voters. And finally, chapter 4 explores the accountability consequences of earmarking tax revenue for particular public expenditures in an online survey experiment in Ghana. It is important to note that the three chapters do not directly leverage variation in state capacity *per se*; they do not examine direct increases in taxation, for example. Instead, the two

taxation chapters are about different ways of taxing society. However, the themes examined in all three chapters represent increases in state capacity in several indirect ways. First, delegating tax collection to city chiefs in the DRC has been shown to increase the extractive capacity of the state, a key component of state capacity (Balán et al., 2022). Second, fighting corruption is widely regarded as one of the key ways to strengthen inclusive governance and thus state strength (Olken and Pande, 2019), as opposed to the alternative where widespread corruption hinders the development of state capacity. Finally, earmarking taxes for particular expenditure categories requires greater state capacity due to the management of several different funds and dedicated public agencies, as opposed to only one general government revenue fund (McCleary, 1991). These themes thus represent a state's attempt to increase its reach vis-à-vis society.

To problematize even further, state capacity as a concept fulfills multiple functions in this thesis. As stated in the previous paragraph, in all chapters, it is, indirectly, the independent variable, either in the form of fiscal capacity or general government effectiveness. In some chapters, state capacity delimits the scope conditions of theory and empirics, such as in chapters 1 and 3 where some state capacity is needed in order to engender the changes in the independent variables that the chapters are leveraging. In yet others, state capacity serves in case selection, such as in chapter 2, where the deliberate selection of one low-capacity and one high-capacity context yields further comparative insights into the workings of anti-corruption political campaigns.

The main insights I document in this thesis are the following. First, drawing on evidence from a field experiment in the DRC, I show that delegating tax collection to local elites such as city chiefs does not impede citizens' ability or willingness to hold both chiefs and the state accountable for policy outcomes. Second, based on conjoint experimental evidence from India and the U.K., I demonstrate that promising to fight corruption as a central policy platform is electorally punished by voters compared to promising to improve the economy. Finally, based on preliminary evidence from an online survey experiment conducted in Ghana, I show that citizens might engage in political life in response to taxation to the same degree, regardless of whether tax revenue is earmarked or not.

Thus, this thesis probes some of the micro-level behavioral assumptions underlying much of the political economy literature on the relationship between society and the state. The findings thus speak to several literatures in the field of comparative politics. First and foremost, the insights offered here are relevant to a literature that studies state-society relations and its implications for state capacity (Migdal, 1988;

Acemoglu and Robinson, 2019; Stasavage, 2020); it asks about the proper balance of power between state and society as well as about the exact sequencing of political development. Two adjacent literatures that this thesis speaks to have implications for state-society relations. The first is the scholarly study of the relationship between taxation and representation (Ross, 2001; Paler, 2013; Weigel, 2020); it asks whether taxation catalyzes citizen engagement with the state as well as under which conditions it does so. The second is the literature on the relationship between political corruption and electoral accountability (Muñoz, Anduiza and Gallego, 2016; De Vries and Solaz, 2017; Klačnja, Lupu and Tucker, 2020); it asks about the causes and consequences of the frequently observed lack of accountability for corruption as well as about conditions under which corruption voting is more or less strong. In what follows, I consider the arguments of each of the strands of literature in turn.

Migdal (1988) studies why some societies come to be organized in a way that some social organizations can resist state capacity building initiatives of their states. He defines state capacity as “[...] the capabilities of states to achieve the kinds of changes in society that their leaders have sought through state planning, policies, and actions” (p. 4). This is reflected in the penetration of society, regulation of social relationships, extraction of resources, and appropriation or use of resources. States seek to establish social control, that is, the “[...] subordination of people’s own inclinations of social behavior or behavior sought by other social organizations in favor of the behavior prescribed by state rules” (p. 22). However, states encounter various forms of resistance to these projects by numerous social organizations such as ethnic groups, social classes, and villages led by chiefs, landlords, rich peasants or clan leaders. In short, “[...] conflict was over who had the right and ability to make rules in that society” (p. 30). His central argument is that state capacity building initiatives have been hindered in many countries by the mobilization of society through social organizations. In Migdal’s matrix of strong/weak state/society, there is no state in history where social control of both state and society is strong; instead, a strong state can only exist with weak society, and strong society can only imply a weak state.

Stasavage (2020) formulates an argument that has similar implications for state-society relations. He distinguishes analytically between ‘early’ and ‘modern’ democracy: in early democracy, a ruler governs jointly with a council or assembly that is made up of members of society independent from the ruler, helping with governance. Crucially, this meant that democracy existed in lieu of a state bureaucracy, precisely because of a lack of knowledge about the productive capacity of the economy. By contrast, modern democracy entails the election of representatives under

universal suffrage and only has episodic, in contrast to continual, participation by citizens. Stasavage's central argument is about the precise sequencing of political development: if a strong state emerges first, democracy is unlikely to follow an increase in state capacity. Or, as he puts it, "[...] just as early democracy existed in lieu of a state, democratization since 1989 has been more likely to survive when the initial power of the central state is weak" (p. 25). According to this line of reasoning, prior experience with a state bureaucracy makes transitioning to modern democracy much less likely. However, if democratic practices emerge before a state bureaucracy, there might be scope to build the state and society jointly, a process in which bureaucracy is not a substitute, but a complement to democracy.

By contrast, Acemoglu and Robinson (2019) contend that under certain circumstances, state and society grow together in their strengths. They study how liberty embodied by civil society can be preserved in the face of growing oppression and capacity building by a central state and define different forms of state-society relations along this two-dimensional continuum of state strength and society strength: under a despotic leviathan, the state is strong relative to society, whereas the opposite is true for the absent leviathan. Crucially, a shackled leviathan is one where both state and society are strong so that the growing state capacity is closely monitored by civil society through accountability and responsiveness. Or, as Acemoglu and Robinson (2019) put it: "It is powerful, but coexists with and listens to a society that is vigilant and willing to get involved in politics and contest power. It is what we'll call a Shackled Leviathan. In the same way that the Leviathan can shackle the Texan gunmen, so that they cannot do harm to ordinary citizens, it can itself be shackled by common people, by norms, and by institutions - in short by society" (p. 27). The central question of the book is how these shackles emerge. The answer is: under conditions of cooperation and trust which are the result of societal control and mobilization, according to the prerogative 'trust and verify'. Through institutionalized and noninstitutionalized involvement of society in politics as well as collective action, the power of society and the power of the state jointly build state capacity to resolve conflicts, control feuds and provide public services while at the same time increasing society's control over it and transforming prevailing norms.

This thesis documents at times sizeable and significant political responses to the state's effort at increasing its capacity, either in the form of delegating taxes or fighting corruption. At other times, for example when earmarking taxes, increasing capacity may have no effect on citizen participation in political life. This suggests greater need for nuance in the precise sequencing of political development: increasing state

capacity does not necessarily lead to autocratic forms of government, and citizens—above and beyond elites—play an active role in shaping state-society relations.

Two adjacent literatures that this thesis speaks to have implications for state-society relations. The first is the scholarly study of the relationship between taxation and representation. How do citizens react to the state’s effort at increasing its fiscal extractive capacity? Ross (2001) studies whether oil wealth has negative effects on democracy, and if, by implication, taxation has positive effects on societal mobilization. The “rentier effect” of oil wealth, for which he finds evidence, is related to the revenues from the sale of oil which allows states to tax society less heavily or not at all, and “[...] the public in turn will be less likely to demand accountability from – and representation in – their government” (p. 332). Paler (2013) then studies the question of whether taxation makes citizens more or less likely to demand accountability from leaders at the micro level. She finds that tax revenue increases monitoring and anti-incumbent political action yet fails to increase participation in politics more generally. The results suggest that this is likely due to an “ownership mechanism” whereby citizens perceive more of the total budget as derived from their taxes and thus feel entitled to certain benefits. Finally, the field experiment reported in Weigel (2020) addresses the same question with slightly different results. Here, the results are evidence for a combination of a “bargaining mechanism” and an “updating mechanism”. According to the former, building fiscal capacity gives citizens a bargaining tool to demand more inclusive governance before complying with paying taxes. According to the latter, increasing fiscal capacity sends a signal to citizens that the state is now more capable of delivering on public goods and services, raising the expected benefits of participation in politics to influence future tax policy and government spending.

This thesis adds to this literature in two ways. First, certain forms of tax collection may not yield the kind of differential participation dividends that the literature would expect: delegating tax collection to local elites does not diminish citizen incentives to hold either chiefs or the government accountable for policy outcomes. Second, by contrast, we should not expect certain forms of taxation to yield higher participation dividends than others: earmarking taxes does not increase citizens’ incentives to participate politically.

The second adjacent literature is a body of work that studies the relationship between political corruption and electoral accountability. How does society react to a decline in state capacity brought about by corrupt practices? Muñoz, Anduiza and Gallego (2016) study why electoral accountability for political corruption has often

been absent from many elections worldwide. They observe that corruption only has limited effects on voting behavior and election results, and this failure to retrospectively sanction misconduct at the ballot box undermines accountability. They find support for an “implicit exchange mechanism” as well as for a “credibility of information mechanism”. According to the former, good performance trumps corruption, and according to the latter, accusations of corrupt practices from opposition parties lack credibility. De Vries and Solaz (2017) review a host of breaking points at which democratic accountability for corruption might discontinue to work. According to their framework, the accountability relationships that are typically assumed for corruption might break down at three stages: information acquisition, blame attribution, and behavioral responses. This then leads to a lack of retrospective punishment of corruption. Furthermore, voter coordination around corruption at the ballot box has often been absent. Finally, Klašnja, Lupu and Tucker (2020) study the conditions under which voters hold corrupt politicians accountable. They observe that in their survey experiments, voters withdraw electoral support from corrupt candidates. However, they also find that information about possibly attractive side benefits of corruption diminishes the degree of electoral accountability for corruption.

This thesis sheds light on questions asked in this literature. Most importantly, electoral accountability for corruption may be weak not because of the lack of credible alternatives or partisan bias. Instead, voter preferences over fighting corruption might be less sensitive to political activation efforts and issue entrepreneurship than the corruption voting literature assumes.

Tangentially, this thesis yields implications for two additional literatures. More narrowly, the results speak to a body of work that concerns the democratic politics of redistribution. This literature has been preoccupied with examining the effect of inequality—mediated through taxation and redistribution—on democratization and voting behavior (Boix and Stokes, 2003; Acemoglu and Robinson, 2005; Pontusson and Rueda, 2010; Anderson and Beramendi, 2012). In this thesis, I explore the possibility that political behavior can powerfully shape the tax state in the first place and thus that part of inequality that is due to differential tax rates and the politics of redistribution. More broadly, my work has implications for the literature on institutions and institutional change. North (1991) has shown that effective formal institutions emerge when the state is strong and the government possesses sufficient coercive capacity. This is especially pertinent in contexts of informal institutions and weak enforcement (Blattman, Hartman and Blair, 2014; Christensen, Hartman and Samii, 2021*b*; Hartman, Blair and Blattman, 2021). Complementing this idea,

the results presented in this thesis show that state capacity can emerge through powerful societal mobilization responses to the state's attempt at increasing capacity. Examining this process at the micro level—which North (1991) would call “a shackling of the arbitrary behavior of the state over economic activity” (p. 109)—is one of the main contributions of this thesis.

Chapter 2

Delegating Tax Collection Does Not Adversely Affect Demand for Accountability: Evidence from an Experiment in the D.R. Congo

(Co-authored with Augustin Bergeron, Laura Paler, Gabriel Tourek, and Jonathan Weigel)

2.1 Introduction

Past scholarship suggests that tax collection enhances citizen political engagement and efforts to hold the government accountable (Ross, 2001; Paler, 2013; Martin, 2016; Weigel, 2020). Reliance on broad-based taxation is also thought to have catalyzed citizen political participation historically in the formation of early modern European states (Tilly, 1992) and remains at the heart of the social contract between citizens and the state in many democracies today (Levi, 1988; North and Weingast, 1989; Timmons, 2005; Prichard, 2015).

Despite considerable evidence about the links between taxation and accountability, little is known about how tax collection by different actors affects accountability. In many developing countries, governments delegate tax collection to local non-state leaders such as chiefs (Boone, 2003; Baldwin, 2015; Jibao, Prichard and van den Boogaard, 2017; Gottlieb and Hollenbach, 2020). Local leaders often enjoy informational advantages relative to state collectors and thus may generate higher tax revenue (Balán et al., 2022). Additionally, research on decentralization suggests that delegating tax collection could improve accountability by bringing state functions

closer to citizens (Faguet, 2004; Mookherjee, 2006). Local elites might simply be more accountable to the population, even in authoritarian settings (Degrave, 2023).

In many states with low capacity today, tax collection is frequently delegated to local actors. Historically, these were mostly tax farmers. In many sub-Saharan African countries today, local actors involved in tax collection are either rebel groups (Sanchez de la Sierra, 2020), local elites (Iversen et al., 2006; Gottlieb and Hollenbach, 2020) or customary and neo-customary chiefs (Mamdani, 1996; Boone, 2003; Baldwin, 2015; Jibao, Prichard and van den Boogaard, 2017; van den Boogaard, 2021; Balán et al., 2022; Bergeron et al., 2022). In all these cases, local agents have played a key role in aiding revenue extraction.

However, delegating tax collection to local non-state leaders could come at the expense of citizen demand for accountability targeted at the state. Specifically, delegating tax collection to chiefs could redirect accountability demands away from the state and toward the chief if it causes citizens to view the chief as having relatively more responsibility in public spending and public goods provision (Bodea and LeBas, 2016; Lust and Rakner, 2018; Henn, 2020). Additionally, increasing the number of actors involved in taxation could make it harder for citizens to know whom to hold accountable for taxing and spending decisions (Healy and Malhotra, 2013; Harding, 2015; Martin and Harper, 2021). Delegating collection could also unwittingly send a signal of weak state revenue and spending capacity and dull citizens' incentives to hold the state accountable (Weigel, 2020). In short, while delegating tax collection could have some advantages, it could also shift the locus of accountability and adversely affect the formation of the fiscal contract between the citizen and the state. This is important because strong citizen-state linkages are thought to engender better development outcomes than reliance on informal institutions (Besley and Persson, 2011).

We examine the accountability consequences of delegating tax collection to chiefs using a randomized policy experiment implemented in Kananga, a provincial capital in the Democratic Republic of the Congo (DRC). In this study, we extend the experimental design described in Balán et al. (2022) in which 367 neighborhoods were randomized into property tax collection by state agents or local chiefs. Until 2018, all tax collection had been conducted by state agents but in that year the government delegated this responsibility to city chiefs in randomly selected neighborhoods. This allows us to study the effects on accountability of delegating tax collection to chiefs (the treatment) relative to the status quo of tax collection by state agents (the control). Our design holds constant all other aspects of taxation—including the level of

government that received the tax revenue (the province)—allowing us to isolate the accountability effects of the state’s decision to use different types of agents to collect taxes. Chiefs are the relevant non-state actor for tax collection and public goods provision in Kananga, as they are elsewhere in sub-Saharan Africa (Baldwin, 2015). While city chiefs in Kananga have closer ties to the state than customary chiefs, they are nonetheless seen as independent, non-state actors by Congolese citizens.

To obtain a real-world measure of demand for accountability after the tax campaign, we partnered with the provincial government’s Division of Social Affairs (DIVAS). In 101 of the original 367 study neighborhoods, 20 percent of households ($N = 2,706$) were given a chance to request a community audit of the chief and/or the state concerning a recent government-run antipoverty program implemented in each neighborhood. We use this design to test two main pre-registered hypotheses: that tax collection by chiefs (relative to tax collection by state agents) reduces the demand for accountability directed at the state (H1); and that chief (relative to state) tax collection enhances accountability demands directed at the chief (H2). Support for both hypotheses would be consistent with the concern that chief collection might enhance chief accountability at the expense of state accountability.

Citizens were informed that neighborhoods that submitted the most requests would have audit meetings organized by a respected local civil society organization. Submitting the request form required costly action on the part of citizens—delivering a completed form to a locked drop box in the city center—and thus provides a real-world measure of citizens’ accountability demands and willingness to take collective action. Crucially, citizens could request audits of the state, the chief, or both, with each request incurring an independent cost.

Our empirical analysis begins by validating the first stage—that tax collection by local chiefs was salient to citizens and effective in extracting revenue. We replicate the results of Balán et al. (2022) in our sample and find that chiefs in fact outperformed state collectors in collecting taxes. We also validate that the distinction between chiefs and state agents as tax collectors was highly salient to property owners. Importantly, a randomized evaluation of a prior city-wide property tax campaign involving only state agents as collectors found substantial treatment effects on citizen demand for accountability, demonstrating the existence of tax-based accountability in Kananga (Weigel, 2020). These results support the appropriateness of investigating the consequences of delegating tax collection to local chiefs on accountability demands in this context.

However, we find no support for the hypothesis that delegating tax collection to local chiefs altered citizens' efforts to hold the chief or the state accountable. Across all neighborhoods, about 17 percent of citizens requested audits of the chief and the state. There is no indication that chief tax collection differentially affected demands for accountability directed at the state or the chief. Additional results employing original survey data corroborate the null results. Exploring our theorized mechanisms in survey data further reinforces the (null) results. We find no evidence that chief taxation reduced the perceived responsibility or capacity of the state, or increased the perceived responsibility or capacity of chiefs. The main implication is that, even in a weak state like the DRC, it is possible to delegate tax collection to local non-state actors like chiefs without undermining the demand for government accountability.

We rule out several alternative explanations for the null results. Specifically, we address concerns that audit meetings did not meaningfully capture citizen demand for accountability; that tax collection was not salient to citizens when deciding whether to submit their audit forms; and that taxation increased the perceived coercive power of chiefs and the state, suppressing demand for accountability. Additionally, we consider two explanations for why chief collection did not increase demand for chief accountability. These include the possibility that chiefs used their informational advantages to target loyal or weak populations for taxation or that tax collection by chiefs increased the perception that they were co-opted by the state—either of which might reduce the will or ability of citizens to try to hold the chiefs accountable. None of these alternative explanations are supported by the data.

All in all, this paper provides (to our knowledge) the first causal evidence of the effect of delegating tax collection to local leaders on accountability and the tax-based social contract. Despite a strong and salient treatment, the field experiment we study provides no evidence that delegating tax collection to local non-state actors undermined accountability towards the state. Our evidence appears most consistent with the claim that this is because citizens remained aware that chiefs remitted revenue to the state and that the state was centrally involved in public goods provision. The main implication is that, even in weak states like the D.R. Congo, it is possible for the state to delegate tax collection to local non-state agents and increase revenue without weakening accountability. It also suggests that states cannot strategically rely on delegation to avoid accountability pressures. We return to a discussion of scope conditions and the external validity of these findings in the conclusion.

This paper proceeds as follows. In section 2.2, we lay out the theoretical framework from which we derive our hypotheses. In section 2.3, we elaborate on the experimental

design and implementation of the randomized controlled trial in Kananga. In section 2.4.1, we present our first-stage results to show that the treatment we evaluate was strong and salient. In section 2.4.2, we discuss our main result that chief tax collection altered citizens' effort to hold the chief or the state accountable. In section 2.4.3, we rule out several alternative explanations for these null results. Section 2.5 concludes and discusses the external validity of our findings.

2.2 Theoretical Framework

2.2.1 Voice in Response to Tax Collection Delegation

Here we develop a conceptual framework for thinking through the accountability consequences of delegating tax collection to local elites. In particular, we model citizens' choice of whether to hold either the government or the chief accountable via political action in response to experiencing property tax collection by city chiefs. We are thus able to derive our two main pre-registered hypotheses.

Society consists of a continuum of citizens. The government announces a property tax campaign, carried out either by local elites or government officials. Crucially, for the purposes of this framework, the government makes this choice exogenously to anticipated citizen bottom-up accountability pressures. This setup gives citizens varying incentives to exert voice to achieve various different goals towards each actor.

In particular, citizens can choose to exert voice or to stay quiet in response to the government's tax campaign. The main goal of exerting voice is to hold a particular actor, elites or the government, accountable for recent policy programs and to demand a citizen meeting. However, there are costs associated with voice (in the form of time and effort and sometimes monetary costs), and the potential benefits from voice need to be discounted by a probability of successful demand-making on leaders. This probability can be thought of as political efficacy and responsiveness of the political system to citizen demand-making. The central question is whether citizens change their political response to the government's tax campaign when local elites are involved in collecting taxes.

Previous literature suggests they do. On the one hand, citizens might demand greater accountability from their government because these decentralization efforts have brought it closer to the people (Faguet, 2004; Mookherjee, 2006). On the other hand, citizens might demand less accountability from their government because the locus of accountability shifts to the chiefs (Bodea and LeBas, 2016; Lust and Rakner, 2018; Henn, 2020). In what follows, we explicate the particular logic of the latter

argument in greater detail to arrive at our two main hypotheses, namely, that chief tax collection reduces demand for accountability directed at the state (H1) and that it increases demand for accountability directed at the chiefs (H2).

Less accountability demand towards the state and more accountability demand towards the chiefs might be produced via several channels. First, citizens could update their beliefs about which actor has relatively more responsibility on public spending and public goods provision. In the context of our conceptual framework, this would mean that the benefits from voice directed towards chiefs would increase in response to chief taxation as the chiefs are thought to be mainly responsible for public finance decision-making. Indeed, recent research shows that there is wide variation in the extent to which traditional authorities assume responsibility for public goods provision. Henn (2020) shows that chiefs are differentially responsible for public goods provision depending on whether they are integrated in the country’s constitution, affecting the degree of complementarity and substitutability with state provision. Lust and Rakner (2018) demonstrate how citizens make substantial contributions to public goods provision via channels other than the state—through traditional authorities, for example, a process they call “social extraction”. Finally, non-state provision of public goods—through, for example, community-provided public goods—has been shown to affect tax compliance and therefore state building (Bodea and LeBas, 2016). This substantial investment in public goods by traditional authorities can be explained by longer time horizons (Baldwin, 2015).

Second, increasing the number of actors involved in taxation could make it harder for citizens to know whom to hold accountable for taxing and spending decisions. In the context of our conceptual framework, this would imply an increase in the costs of voice directed towards the government as it would be increasingly costly for citizens to discern responsibility and thus to direct voice at the state. Indeed, existing literature suggests that it is more difficult for citizens to exert accountability demands the more actors are involved in the process. For example, Healy and Malhotra (2013) show that this process of attribution, “the process by which voters allocate blame and credit to political actors”, is more difficult for voters when multiple actors are involved. In a similar vein, Gomez and Wilson (2001) demonstrate that only politically sophisticated citizens are able to link pocketbook financial considerations to government policy and assign responsibility across multiple political institutions, a process they call “heterogeneous attribution”. Finally, the core argument in Harding (2015) is that government accountability is strong only when citizens are able to correctly attribute

policy responsibility, a process that might be undermined when multiple actors are involved, such as through decentralization or international agencies.

Finally, delegating tax collection could unwittingly send a signal of weak state revenue and spending capacity and dampen citizens' incentives to hold the state accountable for tax decisions. In the context of our conceptual framework, this would equate to a decrease in the benefits from voice directed towards the state as the state is seen as less and less capable of taxing people, thus reducing the benefits derived from holding it accountable over taxation issues. Indeed, recent research demonstrates that when citizens perceive state capacity to be high, demands for accountability are higher too as the expected benefits from participation are higher (Weigel, 2020): "Observing a systematic citywide tax campaign may lead citizens to believe that the government is now more capable of affecting their future well-being by continuing tax collection or by providing more public goods" (p. 41). By contrast, signals of weak state revenue and spending capacity could thus lower bottom-up accountability pressures.

Accountability for policy outcomes can only happen with targeted responsibility attribution, that is, citizens can only hold those institutions accountable who they think are responsible for these outcomes. Because the target of the collective action campaign, an anti-poverty program, is run by the government in both chief and state collection neighborhoods — so that the locus of accountability is technically the same — we note here three additional points that are worth emphasizing. First, the very act of collecting taxes by a given actor might shift citizens' *beliefs* about who is responsible for public goods provision, perhaps because of confusion at the multiple actors involved (Healy and Malhotra, 2013; Harding, 2015; Martin and Harper, 2021), even though technically, it is still the government who bears ultimate responsibility for the anti-poverty program. Second, in many sub-Saharan African contexts, chiefs are responsible for the provision of a wide array of public goods (Baldwin, 2015). This suggests that chiefs matter in service provision and act sometimes independently of the government. In our context, city chiefs are responsible for a public works program organized entirely by them called *salongo* (Balán et al., 2022). Finally, at the same time, the only way to hold chiefs accountable in this context is not through local elections but through petitioning as envisioned in the collective action campaign. City chiefs are nominated by a council of elders in the community but with indefinite and often lifelong tenure (Balán et al., 2022).

To summarize our expectations, we hypothesize the following:

H1 Tax collection by chiefs (relative to tax collection by state agents) reduces the demand for accountability directed at the state.

H2 Tax collection by chiefs (relative to tax collection by state agents) increases the demand for accountability directed at the chief.

2.2.2 Scope Conditions

These bottom-up accountability pressures are only possible in settings in which there is scope for citizens to actively participate in political life. In terms of the scope conditions of our theory, then, our argument primarily applies to settings in which there is scope for political participation to flourish, regardless of whether this applies to the national or subnational level. In order to hold agents to account, there needs to be room for active participation and citizen political efficacy (Lieberman and Zhou, 2021), that is, “a person’s sense of being capable of affecting change” (p. 3). In addition, we need to assume that citizens routinely overcome dilemmas of collective action, either via the provision of private benefits (Olson, 1971), psychological rewards (Bueno de Mesquita and Shadmehr, 2022), or reduced social costs (Paler, Marshall and Atallah, 2018). At the national level, our case, the DRC, does not qualify formally as a democracy (it is classified as “not free” according to Freedom House with a democracy score of 19/100).¹ However, we observe empirically that across all neighborhoods, about 17 percent of citizens requested audits of the chief and the state. This suggests that a sizeable portion of citizens are generally able to overcome collective action problems.

There is a vast amount of subnational variation in the extent to which citizens are active in political life on a regular basis. Lots of activities in Kananga, Kasai Central, can be viewed as a “public transcript” (Scott, 1990), that is, a more formal, individualized and bureaucratic form of demand-making than “hidden transcripts” that are more informal and subtle in nature, such as personal networks or customary institutions (Honig, 2022) that are codified in unofficial rules and norms of, for example, solidary associations and associational life (Tsai, 2007; Wilfahrt, 2022; Harris and Honig, 2023). Qualitative evidence suggest that in our context, this is unlikely to be a source of measurement error. Habits of associationalism and claim-making have deep historical roots Kananga and the DRC more broadly. Nowadays, there is a strongly formalized system of making complaints to the government and other authorities where citizens oftentimes directly voice their concerns in formal meetings

¹<https://freedomhouse.org/country/democratic-republic-congo/freedom-world/2023>.

or by writing letters and petitions, often transmitted through local civil society organizations. This seems even more plausible given the prevalence of strong norms in precolonial times around public debate and contestation of authority and power in our context (Douglas, 1963; Vansina, 1990, 2010) as well as in the wider African context (Bolt et al., 2022).

Our theoretical argument applies to a broad class of local actors that could be classified as “elites”. Historically, tax collection was often delegated to tax farmers. We would expect that the *process* of tax collection, including the collector itself, matters in its own right independent of revenue distribution. Historically, tax farming carried out by local notables incited protests in early modern Europe (Hallenberg, 2009), the Ottoman Empire (Aytekin, 2013), or China (Prazniak, 1980), even when revenue was distributed by the government. There is evidence that the process of tax collection establishes links between citizens and tax collectors directly, e.g. through informal and collusive tax bargaining (Khan, Khwaja and Olken, 2016) or through shaping citizens perceptions of state capacity (Weigel and Ngindu, 2021). In many sub-Saharan African countries today, local actors involved in tax collection are either rebel groups (Sanchez de la Sierra, 2020), local elites (Iversen et al., 2006; Gottlieb and Hollenbach, 2020) or customary and neo-customary chiefs (Mamdani, 1996; Boone, 2003; Baldwin, 2015; Jibao, Prichard and van den Boogaard, 2017; van den Boogaard, 2021; Balán et al., 2022; Bergeron et al., 2022).

Empirically, however, we focus on a narrower set of actors known as city chiefs which are, in particular, different from customary chiefs. City chiefs are local notables who are nominated by a council of elders in the community and approved of by city government officials but with indefinite and often lifelong tenure. In the property tax campaign, each city chief was assigned a local assistant for training and assistance in the campaign, thus forming a collection pair. These chiefs are widely seen as informal political and non-state leaders and traditional authorities (as opposed to formal state agents) whose main responsibilities are to mediate local disputes and to help maintain local infrastructure through the levy of an informal labor tax (Balán et al., 2022; Bergeron et al., 2022). These city chiefs do not receive compensation in their positions but instead rely on assuming other paid positions. Chiefs are thus accountable to both the citizens in their jurisdiction and to the state. Although similar, city chiefs are not equivalent to customary chiefs who are believed to possess and exercise supernatural powers, but are common in many francophone African countries (Boone, 2003; Honig, 2017) with salient roles in the collection of property

taxes (Nguema, 2005; Cogneau et al., 2020).²

2.3 Experimental Design and Implementation

2.3.1 Setting

Our study was embedded in a randomized policy experiment examining a property tax campaign in the city of Kananga, DRC, in 2018. Kananga, the capital of the Kasai Central province, has a population of nearly 1 million inhabitants, and an average monthly household income of \$106 (PPP \$168, see Balán et al. (2022)). The DRC is a low-capacity, “fragile” state, with a tax-to-GDP ratio ranking well below the OECD average.³ Following a large decentralization reform starting in 2006, local governments of the country’s eleven provinces were given either exclusive jurisdiction or shared responsibility with the central government in Kinshasa in many policy domains. Provincial governments can retain 40% of revenue raised within their jurisdictions and levy their own taxes (Englebert and Mungongo, 2016; Paler et al., 2017). The broader impacts on compliance and revenue of this tax experiment, which varied whether tax collectors were state agents or city chiefs, are evaluated in Balán et al. (2022).

2.3.2 Treatments

In the property tax campaign, tax collectors went door to door constructing a previously non-existent property register and providing information about the property tax in 367 neighborhoods in Kananga. During property registration, tax collectors delivered tax letters to the owners containing their tax liability and other information about the property tax. After completing the neighborhood property registry, the tax collectors returned for follow-up tax collection visits. Collectors used portable printers to issue receipts to taxpayers. Both city chiefs and government agents received identical training, followed identical collection procedures during the campaign, and had identical financial incentives to collect taxes.⁴

²We pre-registered our experiment at <https://www.socialscisceregistry.org/trials/4683>. In this pre-analysis plan, several other auxiliary hypotheses are specified. The results of tests of these are reported in Appendix A.5. Furthermore, this study has been approved by the London School of Economics Research Ethics Committee: REC Ref. 000972.

³<https://data.worldbank.org/indicator/GC.TAX.TOTL.GD.ZS>.

⁴See Balán et al. (2022) and Bergeron et al. (2022) for additional details and the full evaluation of this tax campaign.

In this study, we compare the two main treatment arms that vary *which actor* collected property taxes for the state: state agents or city chiefs.⁵ In *state* collection neighborhoods, tax collection was carried out by staff of the provincial tax authority. State collectors were nonsalaried contractors who had undertaken previous work for the tax ministry and other branches of the provincial government. In *chief* collection neighborhoods, tax collection was carried out by city chiefs, who are local notables nominated by elders in the community and approved of by city government officials but with indefinite and often lifelong tenure. They help adjudicate disputes and oversee community-led maintenance of local roads and infrastructure. City chiefs have less power and are more easily monitored by the state than customary chiefs in villages (Baldwin, 2015). Nevertheless, city chieftaincy is an important ‘neo-customary’ institution across Francophone Africa (Boone, 2014), and Congolese citizens view chiefs as non-state agents. These city chiefs do not receive compensation in their positions but instead rely on assuming other paid positions.

As evidence, in our baseline data, 77.8% of a random sample of citizens said they had “a lot” or “some” trust in city chiefs, whereas only 46%, 42%, and 40% said so for the national government, provincial government, and state tax collectors, respectively. The difference between views of chiefs and different state agents—which are each closely clustered together—shows that citizens perceive chiefs as a distinct institution from the state.⁶

In the property tax campaign, each city chief was assigned a local assistant for training and assistance in the campaign, thus forming a collector pair. In particular, both types of collectors worked in pairs which were randomly assigned to two neighborhoods per month with re-randomization every month. The policy experiment had additional treatment groups that were evaluated in other work done in this context (Balán et al., 2022; Bergeron et al., 2022). The additional treatment groups cross-randomized information provision and examined a combination of state and chief collection.

2.3.3 Outcome Measures

The main goal of this paper is to examine the effect of chief tax collection on citizen willingness to hold the state and chiefs accountable. Studying this requires a context

⁵These same experimental arms are referred to, respectively, as *central* and *local* collection in Balán et al. (2022) and Bergeron et al. (2022).

⁶See Balán et al. (2022) and Nzongola-Ntalaja (1975) for more information about the institution of city chieftaincy in Kananga.

where both chiefs and state agents would reasonably be perceived as responsible for public goods provision. We examine accountability demands in the context of a government cash transfer program implemented by the Division of Social Affairs (DIVAS) of the provincial government one year after the tax collection campaign. For this program, city chiefs were instructed to allocate eligibility tickets to the poorest quintile of households in each neighborhood (Bergeron et al., 2022). Because of their local information, chiefs are often involved in targeting scarce state benefits in developing countries (Alatas et al., 2012; Basurto, Dupas and Robinson, 2019).

We collect data on accountability demands targeted towards the state and the chiefs in the context of this program in a random sample of 101 (51 *chief* and 50 *state* collection neighborhoods) of the original 367 study neighborhoods. We also collect data from five neighborhoods assigned to a pure control group which kept the old “declarative” system in which individuals were supposed to pay the property tax in person at the tax ministry. In particular, five neighborhoods in Kananga were assigned to keep the status quo system in which taxpayers had to pay themselves at the tax ministry. In this arm, property registers, assignment of tax IDs and distribution of tax letters were conducted by two agents from the revenue authority, with the letters instructing property owners to pay directly at the agency. The small number of neighborhoods in this arm is a result of evidence from the 2016 tax campaign that tax compliance would approximate zero. Given the small number of pure control neighborhoods, we only present the results in Table A.20.

Prior to the distribution of eligibility tickets, 20 percent of households in the 101 study neighborhoods ($N = 2,706$) were randomly selected to receive informational fliers explaining the cash transfer program, as well as forms allowing them to request a community audit meeting (Table A.1).⁷ Citizens were informed that the government and the chiefs were jointly responsible for the distribution of lottery tickets, and that audit meetings would investigate if the program was properly and fairly implemented and that they would be organized by respected civil society organizations in the ten neighborhoods that accrued the highest shares of forms submitted among those distributed. The civil society organizations, RIAC (Network for Transparency and Anti-Corruption) and SOCICO (the Civil Society of Congo), seek to promote government accountability and transparency. The ten meetings took place as promised in early 2020. Because the demand for accountability was measured prior to the distribution of cash transfers, citizens’ decisions are not conditional on how the

⁷We extracted valid form submissions for 2,631 of the 2,706 households.

program was ultimately implemented. This timing allows us to isolate how state versus chief tax collection shapes accountability demands. If, by contrast, citizens could demand audit meetings after the distribution of cash transfers, we would not be able to disentangle the effects of tax collection from the effects of program implementation, which could itself be impacted by chief versus state tax collection.⁸

Crucially, citizens could request audits of the state, the chief, or both, given that both parties jointly organized the cash transfer program. They received audit request forms for each actor, which they were required to submit in separate locked letter boxes in different parts of the city center. An example of an audit request form is depicted in Figure A.1. Submitting both forms thus required incurring higher costs than submitting one. Overall, submitting audit request forms required costly action—incurring transport costs equivalent to the median daily wage—and provides us with a real-world measure of citizens’ accountability demands. We summarize the intervention and the data collection activities in Table A.2.

Importantly, our accountability exercise does not measure *electoral* accountability. Instead, we investigate more general forms of citizens’ demand-making on leaders, be they official state agents or local elites. In particular, calling citizen meetings is the prevalent mechanism for accountability demands in our local context. Similar meetings are precisely the way citizens of Kananga hold their leaders to account. Traditionally, citizens in Kasai would call meetings—referred to as *tshiota* and/or *l’arbre à palabre*—whenever important matters affecting the community arise, especially concerning leadership. These are rooted in pre-colonial traditions in the region (Douglas, 1963; Vansina, 1990, 2010). Thus, our outcome measure represents a meaningful attempt at capturing local accountability pressures on leaders.⁹

⁸One natural concern is that the target of accountability might be unclear if the distribution program has not happened yet at the time citizens make audit meeting requests. Thus, the link between taxation and spending outcomes might have been broken. However, what our paper focuses on is accountability for the distribution of lottery tickets. In fact, citizens perceived the exercise in this way as is confirmed by quantitative text analysis of citizen meetings and submission forms. There, citizens expressed concern about the fairness of the ticket distribution and inquired about lottery winners and losers. In addition, citizens were also able to make connections between the tax campaign and spending targets in the citizen meetings and submission forms, especially with respect to fighting poverty and the distribution of revenue. Unfortunately, we do not have additional survey outcomes that tap into a tax-and-spend connection. Finally, disentangling the effects of tax collection from the effects of program implementation was given priority over making actual program implementation the locus of accountability demands.

⁹One way to examine the possibility of state-centred accountability through elections further would be through turnout data. While we do have self-reported turnout for the 2018 general election, it is unlikely that we would detect any differences between treatment groups due to the closeness in timings of the election and the tax campaign treatment. The tax campaign ran from May to December 2018 and the general election was held in December 2018.

2.3.4 Analysis

We estimate the intent-to-treat effect of *chief* (relative to *state*) tax collection on demands for accountability targeting the chiefs and the provincial government. Our estimation strategy accounts for the fact that randomization was implemented at the neighborhood level using individual-level outcomes and stratified on a combination of geographic location and tax compliance in a previous tax campaign (see Appendix A.1.1). The main theoretical estimand we are interested in is the difference in submission of a citizen meeting request form if a citizen had been exposed to tax collection by the local chief versus if that citizen had been exposed to tax collection by the provincial government. Straightforwardly, our empirical estimand is the Intent-to-Treat (ITT) effect of chief tax collection in the 2018 property tax campaign on submission of a citizen meeting request form. To recover this empirical estimand from the experimental data, our main estimation takes into account the block-cluster nature of the experimental design and is as follows:

$$Y_{ijk} = \beta CHIEF_{jk} + \alpha_k + \mathbf{X}_{ijk}\Gamma + \mathbf{X}_{jk}\Lambda + \epsilon_{ijk}, \quad (2.1)$$

where $CHIEF_{jk}$ is an indicator for whether tax collection in neighborhood k was done by the city chiefs (coded 1) versus by state agents (coded 0). The comparison group consists of neighborhoods where tax collection was carried out by state agents. Y_{ijk} are individual-level measures of submitting a form requesting an audit of the chief or provincial government for citizen i in neighborhood k and stratum j . Additionally, in some analysis below, Y_{ijk} are secondary outcomes from endline surveys conducted 3-6 months and 6-12 months after the tax intervention. All analysis employing survey data draws on all respondents in our study neighborhoods, not just those who received fliers. α_k are fixed effects for the randomization strata, and $\mathbf{X}_{ijk}\Gamma$ and $\mathbf{X}_{jk}\Lambda$ refer to individual-level and neighborhood-level controls (see Appendix Tables A.6 and A.7 for details). Post-hoc power analysis reveals that given 50 clusters per treatment arm and an empirically calculated intra-cluster correlation of 0.1, our design is powered to detect differences in submission rates of 8.5 percentage points, i.e., from a proportion of form submissions of 0.17 in the control group to 0.255 in the treatment group.¹⁰

Treatment groups were balanced on covariates (Appendix A.1.3). We cluster heteroskedasticity-robust standard errors at the neighborhood level. Our main coefficient of interest, β , captures the effect of chief tax collection relative to state tax

¹⁰This power calculation was carried out using EGAP's power calculator at <https://egap.shinyapps.io/power-app/>.

collection. We interpret $\beta < 0$ as support for H1 when Y_{ijk} designates submission of a form requesting an audit of the state and $\beta > 0$ as support for H2 when Y_{ijk} designates submission of a form requesting an audit of the chief. This procedure implements an estimator calculating the individual treatment effect τ within strata and averaging over blocks with weights proportional to block size. We estimate heterogeneous treatment effects via interactions of the respective moderator with the treatment indicator:

$$Y_{ijk} = \beta CHIEF_{jk} * \gamma \mathbf{Z} + \alpha_k + \mathbf{X}_{ijk}\Gamma + \mathbf{X}_{jk}\Lambda + \epsilon_{ijk}, \quad (2.2)$$

where $\mathbf{Z}$ is a vector of various moderator variables which can either be at the neighborhood level k or the individual level i .

2.4 Results

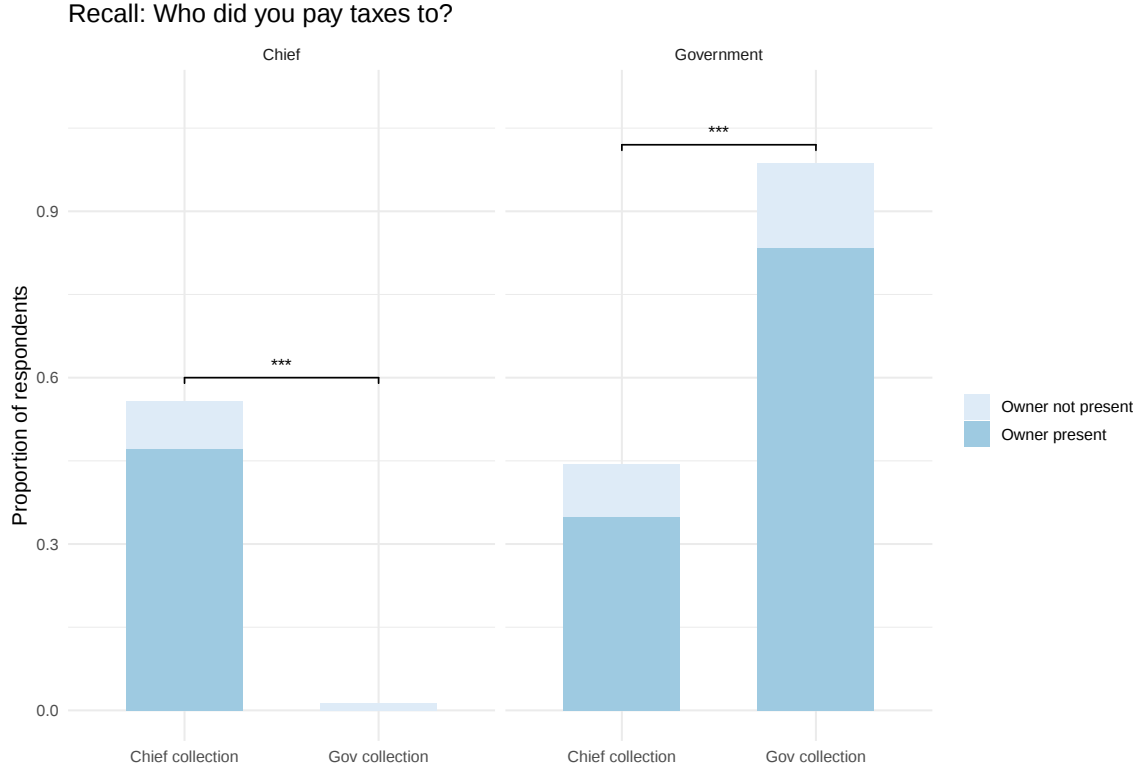
2.4.1 First-Stage Results

We begin by documenting that tax collection by local chiefs was salient to citizens and effective in extracting revenue. During the tax campaign, the average household was visited by tax collectors 1.7 times and an adult was present for 78.9 percent of these visits. Citizens were thus, not surprisingly, highly aware of who collected taxes in the neighborhood (Figure 2.1).¹¹ In state collection neighborhoods, 99 percent of households reported that government agents were in charge of tax collection and only one percent of households thought chiefs were in charge. In chief collection neighborhoods, 56 percent were aware that they paid taxes to the chief. This large difference between treatment and control confirms the relevance of our experimental context for testing H1 and H2. That said, a sizeable minority in chief tax collection places—43 percent—viewed taxes as being paid to the government. This likely indicates that citizens were aware that chiefs were remitting tax revenue to the state and we return to this in a discussion of the null results below.

Further, both types of collectors succeeded at raising revenue (Figure 2.2). We replicate the results of Balán et al. (2022) in our sample of 101 collective action

¹¹One concern is that by asking “Whom did you household pay?” we are not able to differentiate whether respondents take this to mean either who collected the tax (either the chief or the government) or who eventually received the tax (the government). However, the way the endline survey was structured and framed suggests a strong inclination towards the former possibility, that is, the question asked about who collected the tax. Immediately before the payment question, citizens were asked whether their household had been visited in the 2018 property tax campaign, and how many times someone came to sensitize or collect.

Figure 2.1: Salience of Tax Collection by City Chiefs



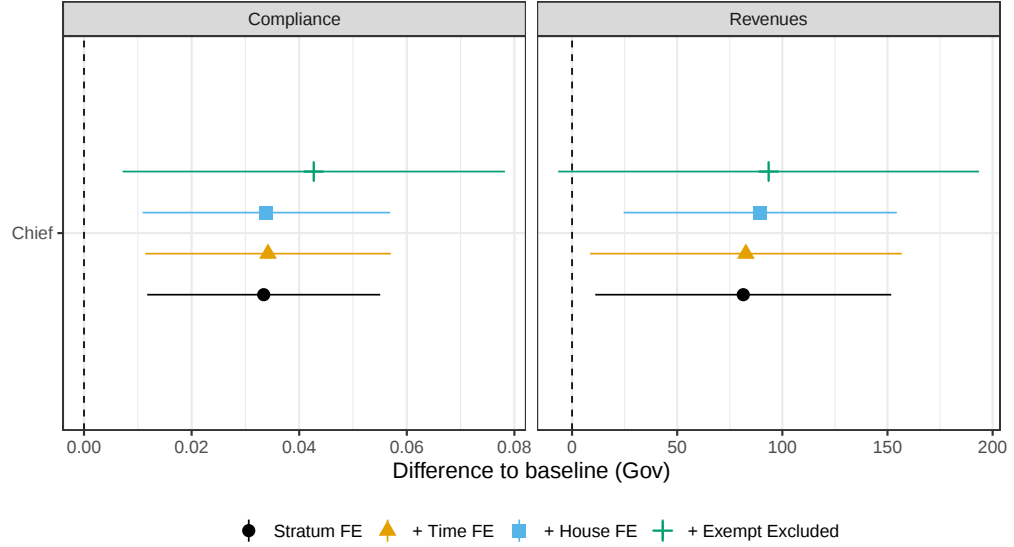
Notes: This figure compares the proportion of respondents who recalled that the Chief (left panel) or the government (right panel) collected taxes across neighborhoods assigned to chief or state collection. We differentiate in each case whether the household owner was present at the time of collection (dark blue bar) or not (light blue bar). P-values are based on cluster-robust covariance matrix estimation (where clusters are neighborhoods) and test for the difference in proportions between state and chief neighborhoods via a regression of the outcome variable on a treatment indicator, including fixed effects for randomization strata.

neighborhoods by comparing compliance and revenue in the *chief* and *state* treatment groups. Across specifications, tax compliance in chief collection neighborhoods was 3.5 percentage points higher, on average, than in state collection neighborhoods (see Balán et al., 2022 for more details). All in all, chiefs and state collectors appear to have executed their duties as the government expected, visiting and collecting taxes from households in the assigned locations.

2.4.2 Main Results

Despite these strong first stage results, we find no evidence that taxation by chiefs is associated with different levels of engagement with the chief or the state. Across

Figure 2.2: Effect of Chief Collection on Tax Compliance and Revenue



Notes: This figures presents the effect of chief tax collection on tax compliance in the 101 collective action neighborhoods. The coefficients come from regressions of the outcome variable (tax compliance for the left panel and tax revenue for the right panel) on a treatment indicator and fixed effects for randomization strata, progressively adding time fixed effects (in yellow), house fixed effects (in blue) as well as restricting the sample to non-exempt properties (green). Standard errors are cluster robust at the neighborhood level and are the basis for the 95% confidence intervals presented.

all neighborhoods, about 17 percent of citizens requested audits of the chief *and* of the state (see Table A.12). However, we find no evidence that tax collection by local chiefs resulted in weaker demand for state accountability (H1). Citizens were equally likely to request audit meetings targeting the state where chiefs and state agents collected taxes (Table 2.1, Columns 1-2). They were also equally likely to request audit meetings targeting the chief (Table 2.1, Columns 3-4). In all cases, the treatment coefficients approximate zero, indicating a genuine null result. We also confirm that these null results are not due to a lack of statistical power. Post-hoc power analysis reveals that given 50 clusters per treatment arm and an empirically calculated intra-cluster correlation of 0.1, our design is powered to detect differences in submission rates of 8.5 percentage points, i.e., from a proportion of form submissions of 0.17 in the control group to 0.255 in the treatment group. The 2016 property tax campaign in Kananga increased citizen participation by about 5 percentage points (Weigel, 2020).

We also examine the effect of chief tax collection on submitting a form demanding accountability by either actor (Table 2.1, Columns 5-6). There is evidence that citi-

zens understood the forms as individual opportunities to hold each actor accountable. Recall that they had two different forms to submit and they had to visit two separate letter boxes to submit them. The ‘either’ indicator captures whether chief collection led to an increase in form submission demanding either state *or* chief accountability. This captures whether citizens submitted one of the two forms indiscriminately, either because they did not distinguish between the two actors or because they viewed them as substitutes for one another.

We also examine the effect on submitting a form for an audit meeting for both actors, that is, whether chief taxation increases the desire to hold both the chief and state jointly accountable (Table 2.1, Columns 7-8). Across treatments, citizens are again equally likely to request audit meetings of either actor or of both actors. These results are robust to a variety of alternative specifications and robustness tests: in Table A.6, we add two different sets of control variables that go beyond neighborhood-level measures and capture individual-level pre-treatment characteristics, and in Table A.7, we use fixed effects for alternative randomization strata based purely on geographical location, not including previous tax compliance. In both cases, we confirm that chief tax collection had no significant and meaningful effect on demands for state or chief meetings.

Table 2.1: Effects of Chief Tax Collection on Demand for Citizen Meetings

	Gov meeting		Chief meeting		Either		Both	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Chief tax collection	-0.014 (0.029)	-0.007 (0.028)	-0.010 (0.030)	-0.003 (0.029)	-0.008 (0.030)	-0.000 (0.029)	-0.017 (0.029)	-0.009 (0.028)
Control mean	0.169	0.169	0.171	0.171	0.180	0.180	0.161	0.161
Observations	2631	2629	2631	2629	2631	2629	2631	2629
Clusters	101	101	101	101	101	101	101	101
Stratum FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Set I	No	Set I	No	Set I	No	Set I

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Coefficients on *Chief tax collection* are OLS estimates with robust standard errors clustered at the polygon level, the unit of randomization and of analysis. *Chief Meeting* is the proportion of citizens requesting a meeting with the chief in a given polygon. *Gov Meeting* is the proportion of citizens requesting a meeting with the provincial government in a given polygon. *Either* is the proportion of citizens in a given polygon requesting either a meeting with the chief or the provincial government. *Both* is the proportion of citizens in a given polygon requesting both a meeting with the chief and with the provincial government. *Chief tax collection* is a binary indicator for whether the polygon experienced tax collection by the chief (1) or by the provincial government (0; control). Here we compare polygons that had a collective action opportunity ($j = 101$) and where either tax collection was done exclusively by the chief ($j = 51$) or by the provincial government ($j = 50$). *Stratum FE* refer to the stratum used for randomization which is a combination of geographic location of the neighborhood as well as previous tax compliance. *Set I* controls include house type, wave of collective action campaign, distances to letter boxes and polygon-level averages of participation and per capita tax revenue in the 2016 campaign.

These null results are confirmed by several additional analyses. First, we explore

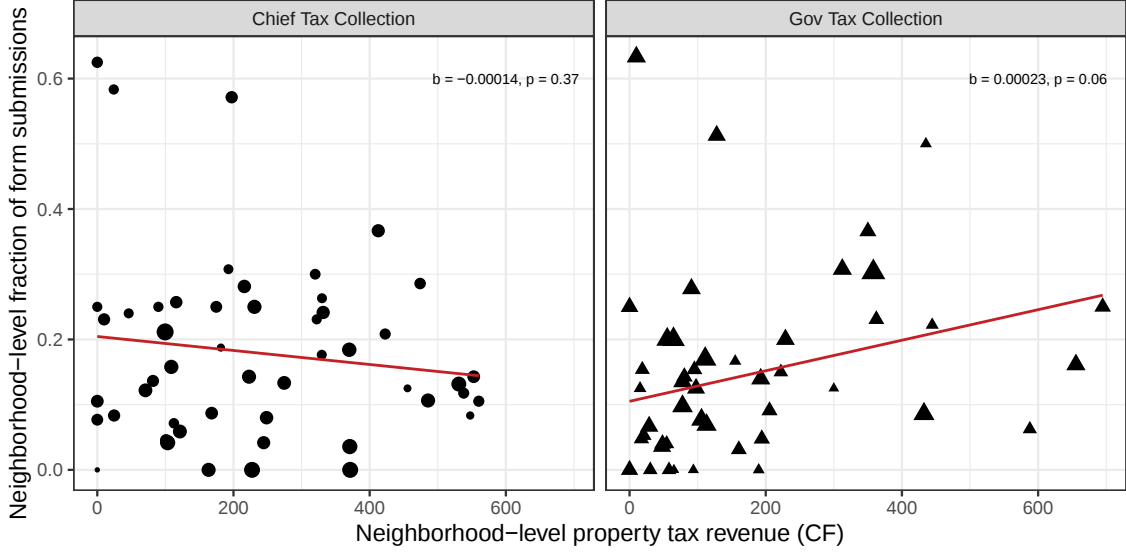
survey-based measures of accountability demands on the government and other local actors (Figure A.3) and political participation, including involvement in party meetings, community meetings, or civic association meetings (Table A.8). In Table A.8, we measure citizens’ self-reported recall of having attended either a community, CSO or political party meeting. Citizens appear to participate roughly equally in these ways across treatment and control as we find no significant differences in response rates between chief collection neighborhoods and state collection neighborhoods.

Second, we examine accountability demands per amount of revenue raised. A natural question is whether the similar level of citizen accountability demand that we observe across treatments in fact reflects an *erosion* of demand in chief tax neighborhoods given that chiefs raised more revenue as documented in Balán et al. (2022). Thus, a natural question is whether we should interpret accountability demands relative to the amount of tax that chiefs and state agents collected.¹² Chiefs raised more revenue than state collectors, so in a sense the ‘participation dividend’ of taxation must be lower for them than it is for state agents. While we do observe a positive relationship between neighborhood-level revenue and participation in the state collection group and a negative relationship in the chief collection group, this difference is not statistically significant (Figure 2.3 and Table A.13). Further to this point, Figure 2.3 plots neighborhood-level property tax revenue against the neighborhood-level fraction of form submissions for *state* neighborhoods (left panel) and *chief* neighborhoods (right panel). The figure’s slopes represent the revenue elasticity of participation in December 2018 US-\$ in both treatments and are close to zero and insignificant. In other words, neither treatment affected the relationship between tax revenue collected and accountability significantly, and the difference between these slopes is not statistically significant at conventional levels ($p = 0.14$). This null result is confirmed when interacting the treatment with the amount of tax revenue collected in the neighborhood as shown in Appendix Table A.13. Ultimately, none of the evidence suggests that chief taxation undermines state accountability or increases chief accountability.

With these slopes in hand, we can calculate the revenue elasticity of participation in December 2018 US-\$ (the time when the tax campaign concluded and all tax revenue was collected), that is, the participation dividend per tax dollar. Moving from \$0 to \$1 would have meant, in December 2018 CF, an increase from CF0 to CF1565.

¹²Not all models of tax-based accountability would motivate this empirical test. If tax *payment*, per se, is thought to generate demands for accountability (e.g., Martin, 2016), then this is a clear testable implication of the model. On the other hand, if being *solicited* to pay taxes, but not necessarily paying, is thought to generate accountability demands (e.g., Weigel, 2020), there would be no predicted correlation between accountability demands and the amount of tax paid.

Figure 2.3: Relationship Between Tax Revenue and Form Submission in State and Chief Neighborhoods



Notes: This figure shows the relationship between tax revenues and engagement (form submission) in Chief (left Panel) and Gov neighborhoods (right panel) in the 101 collective action neighborhoods. For each set of neighborhoods, we plot the neighborhood-level property tax revenue (in Congolese Francs) against the neighborhood-level fraction of form submissions for the government (for Gov neighborhoods) or for the chief (for Chief neighborhoods). We exclude two outlier neighborhoods that exhibited revenue greater than 1,000CF. Shape sizes are proportional to neighborhood sizes. b is the revenue elasticity of participation, i.e., the increase in participation for a 1CF increase in tax revenue.

Multiplying this increase by the slopes b , we get an increase in form submissions by 26.6 percentage points for each tax dollar collected in *state* neighborhoods and a reduction by 21.9 percentage points in form submissions for each tax dollar collected in *chief* neighborhoods. The difference between these slopes, however, is not statistically significant at conventional levels ($p = 0.14$).

The null results are further confirmed by a lack of support for the theorized mechanisms. The central concern behind H1 was that chief taxation would undermine the perceived revenue and spending capacity of the state. Table 2.2 reports self-reported measures of (i) perceptions of the state's or the chief's capacity to fix a damaged road in a given amount of days and (ii) expectations that paying taxes entitles citizens to services distributed by the chief and the government, respectively. Yet we observe no differences across treatments in the perceived capacity of the state, nor do we find any differences across treatments in expectations of benefits from either the state or

the chief.

Table 2.2: Effects of Chief Tax Collection on Expectations about Chiefs and Government

	Expectation of benefits from				Perceptions of capacity about			
	Gov	Gov	Chief	Chief	Gov	Gov	Chief	Chief
Chief tax collection	0.029 (0.117)	0.074 (0.120)	0.013 (0.102)	-0.034 (0.103)	-0.043 (0.055)	-0.050 (0.056)	-0.079 (0.082)	-0.113 (0.088)
Control mean	-0.008	-0.008	0.001	0.001	0.049	0.049	0.038	0.038
Observations	1236	1236	1216	1216	1157	1157	1155	1155
Clusters	101	101	101	101	101	101	101	101
Stratum FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Set I	No	Set I	No	Set I	No	Set I

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Coefficients on *Chief tax collection* are OLS estimates with robust standard errors clustered at the polygon level, the unit of randomization and of analysis. *Chief Meeting* is the proportion of citizens requesting a meeting with the chief in a given polygon. Outcomes come from the Endline II survey. All outcomes have been z-transformed such that coefficients represent standardized mean differences. *Expectation of benefits* measures agreement (5 = strongly agree) with the statement that paying taxes entitles citizens to services by the chief and by the government, respectively. *Perceptions of capacity* measures citizens beliefs about how long it will take the chief or the government to fix a damaged road in days. *Chief tax collection* is a binary indicator for whether the polygon experienced tax collection by the chief (1) or by the provincial government (0; control). Here we compare polygons that had a collective action opportunity ($j = 101$) and where either tax collection was done exclusively by the chief ($j = 51$) or by the provincial government ($j = 50$). We restrict the sample here further to exclude outliers on the capacity questions as well as villas. *Stratum FE* refer to the stratum used for randomization which is a combination of geographic location of the neighborhood as well as previous tax compliance. *Set I* controls include house type, wave of collective action campaign, distances to letter boxes and polygon-level averages of participation and per capita tax revenue in 2016.

Additionally, the concern motivating H2 was that citizens would come to see the chief as having more responsibility and influence in determining public spending and public goods provision, causing some citizens to orient their accountability demands away from the state and towards the chief. But chief tax collection does not appear to have altered citizens' perceived responsibility for providing a range of public goods (Figure A.4). Citizens overwhelmingly viewed the national and provincial governments as primary public goods providers, regardless of which actor collected taxes. Citizens were also aware that ultimately, the tax revenue collected from the property tax campaign is remitted to the state (Figure A.2). It is evident that in both treatment groups, priors expressed in the baseline survey are shifted towards greater responsibility attribution to the national government expressed in the endline survey after the intervention. This suggests that the *chief* treatment did not reduce perceptions of the government's responsibility for spending.

All in all, the results suggest that, in a weak state like the D.R. Congo, the government can delegate tax collection to local non-state actors and enhance revenue collection without weakening the demand for state accountability. It is possible that delegation to chiefs can be revenue-enhancing without having adverse accountability

consequences for the state. This is likely to be true under certain conditions that we discuss more in the conclusion.

2.4.3 Alternative Explanations

To support the substantive interpretation of our null results, we rule out a number of alternative explanations.

Audit meetings and demand for accountability. One possibility is that citizens might not have viewed requesting and attending a meeting as a meaningful way to hold chiefs or the state accountable. This could be the case if, for instance, citizens typically use other means to hold chiefs accountable (Scott, 1990; Tsai, 2007; Honig, 2022; Wilfahrt, 2022; Harris and Honig, 2023). However, similar meetings are precisely the way citizens of Kananga hold their leaders to account. Traditionally, citizens in Kasai would call meetings—referred to as *tshiota* and/or *l’arbre à palabre*—whenever important matters affected the community arise, especially concerning leadership. These are rooted in pre-colonial traditions in the region (Douglas, 1963; Vansina, 1990, 2010).

Other evidence reinforces that the meetings captured meaningful accountability demands. First, respondents who submitted audit request forms were more likely to participate in the audit meetings, suggesting a genuine interest in the meeting itself (Table A.10). Second, citizens’ comments at the meetings and complaint forms they filled out—in both *state* and *chief* neighborhoods—focused on transparency, public integrity, and accountability of the state or the chief in administering the public goods provision program (see the text analysis in Table A.9 and Figures A.6-A.7). Third, survey evidence shows that citizens perceived the audit meetings as a meaningful tool to assert accountability pressures and that the relevant actor will be sanctioned before the meetings took place in both treatment groups. In *chief* neighborhoods, 50.5 percent of survey respondents said it is either likely or very likely that the chief would be sanctioned should a meeting take place, with nearly identical (49.4%) responses in *state* neighborhoods. Finally, Table 2.4 also suggests that baseline levels of fears about making a comment during a citizen meeting with the chief are low and that the treatment does not increase these fears.

Tax collection salience. A second possible explanation for the null results is that tax collection was no longer salient to citizens when they decided whether to demand an accountability meeting. We intentionally avoided an explicit link between the cash transfer program and taxation to avoid mechanical effects on participation (i.e., citizens participating to ask factual questions about taxation rather than wanting

to hold the chief or state accountable for public goods provision). We also measured accountability on average a year after tax collection to avoid capturing short-term effects. That said, it is possible that a lack of salience of taxation at the time of participation explains the similar results across collector treatments.

To examine this, we investigate whether treatment effects were more pronounced in neighborhoods with less time separating the property tax campaign and the collective action opportunity at the neighborhood level. This time gap was random since the timing of both the tax campaign and the cash transfer program were both randomly assigned at the neighborhood level. Audit request form submission does not appear to be higher in neighborhoods where fewer months separated the tax campaign and submission of audit request forms (Table 2.3). Additionally, we conducted a priming experiment embedded in the invitation to return an audit form to check whether reminding citizens to think of themselves as taxpayers strengthened the link between taxation and accountability. However, we find no evidence that respondents addressed as “taxpayers” were more likely to submit an audit request form than respondents addressed as “citizens” (Table A.11), suggesting a lack of salience is not the issue.

Table 2.3: Effects of Tax Campaign Timing on Demand for Citizen Meetings

	Gov meeting		Chief meeting		Either		Both	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Month of Tax Collection	-0.008 (0.005)	-0.008 (0.006)	-0.009 (0.006)	-0.010* (0.006)	-0.009 (0.006)	-0.010* (0.006)	-0.008 (0.005)	-0.008 (0.006)
Control mean	0.141	0.141	0.160	0.160	0.163	0.163	0.137	0.137
Observations	2631	2629	2631	2629	2631	2629	2631	2629
Clusters	101	101	101	101	101	101	101	101
Stratum FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Set I	No	Set I	No	Set I	No	Set I

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Coefficients on *Chief tax collection* are OLS estimates with robust standard errors clustered at the polygon level, the unit of randomization and of analysis. *Chief Meeting* is the proportion of citizens requesting a meeting with the chief in a given polygon. *Gov Meeting* is the proportion of citizens requesting a meeting with the provincial government in a given polygon. *Either* is the proportion of citizens in a given polygon requesting either a meeting with the chief or the provincial government. *Both* is the proportion of citizens in a given polygon requesting both a meeting with the chief and with the provincial government. *Chief tax collection* is a binary indicator for whether the polygon experienced tax collection by the chief (1) or by the provincial government (0; control). Here we compare polygons that had a collective action opportunity ($j = 101$) and where either tax collection was done exclusively by the chief ($j = 51$) or by the provincial government ($j = 50$). *Stratum FE* refer to the stratum used for randomization which is a combination of geographic location of the neighborhood as well as previous tax compliance. *Set I* controls include house type, wave of collective action campaign, distances to letter boxes and polygon-level averages of participation and per capita tax revenue in the 2016 campaign.

Coercive power. We also consider whether tax collection by either the state or chief increased the perceived coercive power of either actor, making citizens less willing to hold either accountable for fear of possible retribution. While the forms were anonymous, it is possible that citizens thought that the chiefs or the government

nevertheless had ways to know who submitted an audit request form. However, according to survey data on the fear of coercion by the chief and the government, there are no differences across treatment and control (Table 2.4, Columns 3-6).

Table 2.4: Effects of Chief Tax Collection on Fear of Coercion and Chief Cooptation

	Chief cooptation		Fear of coercion chief		Fear of coercion gov	
	(1)	(2)	(3)	(4)	(5)	(6)
Chief tax collection	-0.068 (0.073)	-0.042 (0.074)	0.034 (0.121)	0.031 (0.123)	-0.035 (0.128)	-0.049 (0.129)
Control mean	0.032	0.032	-0.016	-0.016	0.023	0.023
Observations	1221	1221	1181	1181	1188	1188
Clusters	101	101	101	101	101	101
Stratum FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Set I	No	Set I	No	Set I
Sample	C + L	C + L	C + L	C + L	C + L	C + L

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Coefficients on *Chief tax collection* are OLS estimates with robust standard errors clustered at the polygon level, the unit of randomization and of analysis. *Chief Meeting* is the proportion of citizens requesting a meeting with the chief in a given polygon. Outcomes come from the Endline II survey. Outcomes come from the Endline II survey and have been z-transformed such that coefficients represent standardized mean differences. *Chief cooptation* measures citizens' perceptions of whether a project undertaken in the neighborhood would be implemented according to the will of the people or that of the government. *Fear of coercion* measures citizens nervousness about making a comment during a citizen meeting with either the chief or the government. *Chief tax collection* is a binary indicator for whether the polygon experienced tax collection by the chief (1) or by the provincial government (0; control). Here we compare polygons that had a collective action opportunity ($j = 101$) and where either tax collection was done exclusively by the chief ($j = 51$) or by the provincial government ($j = 50$). We restrict the sample here further to exclude compounds classified as villas. *Stratum FE* refer to the stratum used for randomization which is a combination of geographic location of the neighborhood as well as previous tax compliance. *Set I* controls include house type, wave of collective action campaign, distances to letter boxes and polygon-level averages of participation and per capita tax revenue in 2016.

Another possibility is that interference could decrease accountability pressures on both actors equally if both the chiefs and the government are feared for their coercive power. While we see some evidence of interference, this does not seem to translate into a positive and significant effect of chief tax collection on perceptions of *government* coercion. As we report in Table A.5, while chief tax collection significantly predicts reporting to have heard about the chief collecting taxes in the neighborhood, 25.9% in *state* neighborhoods also report having heard about chief tax collection. Similarly, around 60% in *chief* neighborhoods had heard about government officials working as tax collectors anywhere in Kananga.

Chief tax collection targeting. Another possible explanation is that chiefs used their informational advantage to collect taxes from loyal citizens such as co-ethnics or clients whom they suspected would pay without increasing accountability demands. This possibility is consistent with the idea that local chiefs could be more

successful at raising taxes from coethnics (Kasara, 2007) or at otherwise suppressing citizen accountability demands (Gottlieb and Hollenbach, 2020). However, there is no evidence that coethnics of the chief were more likely to be targeted in chief relative to state neighborhoods (Balán et al., 2022) or that chief tax collection made ethnicity more salient in neighborhoods (Table A.14). Moreover, coethnics of the chief were no less responsive to treatment in demanding audit meetings targeting the chief or the government (Table A.15).

Chief co-optation. Finally, chief taxation might have lead citizens to believe that chiefs were co-opted by the government and thus less accountable to citizens. To test this possibility, we examine a survey question in which citizens are asked whether they perceive the chief to act more closely in accordance with the preferences of the community or with the preferences of the government in instances where they conflict. According to this measure, delegating tax collection to chiefs did not change citizens’ perception that chiefs are co-opted by the government (Table 2.4, Columns 1-2). In our survey, 26 percent of respondents think the chief is responsive to the state, and 10 percent think the chief is responsive to citizen preferences.

Does Chief Collection Change *Who* Demands Accountability? Finally, we consider the possibility that chief collection does not change the proportion of people who demand accountability but rather the composition or identities of people who demand accountability. While we did not find evidence of differential targeting, and consequently accountability demands, from coethnics, we also check socio-economic status. This arises from the finding in Balán et al. (2022) that chiefs more efficiently target households with higher payment propensity. Standard theory would suggest that we might see chief collection lead to more accountability demand, although it is not clear this would be the case in the targeting of a pro-poor program. We can explore this by looking at predicted ease of tax payment. This variable is used in Balán et al. (2022) to show that state collectors who were advised by city chiefs were more likely to visit and collect taxes from households recommended by chiefs, and emerged from the observation that chiefs must have private information that is useful for targeting visits during the tax campaign. Balán et al. (2022) conclude the following: “Among those visited after registration, a one point increase in the chief’s ability-to-pay (willingness-to-pay) ranking is associated with an 8.3 (5.8) percentage point increase in the probability of payment” (p. 784). However, these recommendations were made only in an additional treatment group of neighborhoods where households were visited by both state agents and the chief. In order to facilitate this analysis in *chief* and *state* neighborhoods, we use a prediction approach based on propensity

scores where we predict properties that chiefs would have recommended, based on a set of household characteristics such as the property owner’s age, gender, employment status, salary, government job status, and ethnic group. These predictions are then transformed into a 1-3 rank. Table 2.5 reports the heterogeneous treatment effects of chief collection by predicted ease of payment. We find that chief tax collection engendered more accountability demands on both the state and the chief when payment propensity was higher. Thus, chief collection elicits more bottom-up accountability pressures from those who had higher predicted payment propensity. We do not find evidence of heterogeneous treatment effects with respect to co-ethnicity or gender (Tables A.15 and A.16). Another possibility related to these findings is that this represents a network effect whereby citizens close to chiefs were more likely to come to citizen meetings regarding the chief. Due to difficulties in measuring closeness in networks, we do not examine this possibility further.

Table 2.5: Effects of Chief Tax Collection: Payment Propensity

	Gov meeting		Chief meeting		Either		Both	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Chief tax collection	-0.123** (0.051)	-0.090* (0.046)	-0.106** (0.052)	-0.078* (0.046)	-0.112** (0.053)	-0.081* (0.047)	-0.116** (0.051)	-0.087* (0.045)
Ease of pay	-0.070** (0.033)	-0.063** (0.029)	-0.088** (0.037)	-0.081** (0.033)	-0.080** (0.036)	-0.072** (0.031)	-0.078** (0.036)	-0.072** (0.033)
Chief X Ease of pay	0.104** (0.047)	0.093* (0.047)	0.097** (0.048)	0.087* (0.048)	0.100** (0.048)	0.088* (0.048)	0.101** (0.048)	0.092* (0.048)
Control mean	0.169	0.169	0.171	0.171	0.180	0.180	0.161	0.161
Observations	1017	1017	1017	1017	1017	1017	1017	1017
Clusters	78	78	78	78	78	78	78	78
Stratum FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Set I	No	Set I	No	Set I	No	Set I

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Coefficients on *Chief tax collection* are OLS estimates with robust standard errors clustered at the polygon level, the unit of randomization and of analysis. *Chief Meeting* is the proportion of citizens requesting a meeting with the chief in a given polygon. *Gov Meeting* is the proportion of citizens requesting a meeting with the provincial government in a given polygon. *Either* is the proportion of citizens in a given polygon requesting either a meeting with the chief or the provincial government. *Both* is the proportion of citizens in a given polygon requesting both a meeting with the chief and with the provincial government. *Chief tax collection* is a binary indicator for whether the polygon experienced tax collection by the chief (1) or by the provincial government (0; control). *Ease of payment* is a categorical variable that represents the predicted tax payment propensity of each compound, with higher values representing higher payment propensities. Here we compare polygons that had a collective action opportunity ($j = 101$) and where either tax collection was done exclusively by the chief ($j = 51$) or by the provincial government ($j = 50$). *Stratum FE* refer to the stratum used for randomization which is a combination of geographic location of the neighborhood as well as previous tax compliance. *Set I* controls include house type, wave of collective action campaign, distances to letter boxes and polygon-level averages of participation and per capita tax revenue 2016.

2.5 Discussion

Using a randomized field experiment in D.R. Congo, we find that the delegation of tax collection from the state to local chiefs does not undermine citizen accountability demands on the state, or increase accountability demands on the chief at the expense

of the state. These null results provide evidence that, in a weak state like the D.R. Congo, delegating tax collection to local non-state actors can raise revenue without incurring an accountability penalty and impeding the formation of a citizen-state fiscal contract. They also suggest that the strategic delegation of tax collection to *avoid* accountability pressure might not be straightforward.

We would, however, note several scope conditions for the interpretation of these findings. First, city chiefs are not customary chiefs. Although they share many characteristics, they have less power and are more easily monitored by the state than customary chiefs in villages. While our results are likely relevant across much of urban sub-Saharan Africa, they should not be generalized to the state delegating tax responsibilities to customary chiefs. Second, more generally, we believe that citizens' awareness that revenue was remitted to the state and the state retained authority in spending and public goods provision is likely a necessary condition for delegating tax collection to have no adverse accountability effects. The results might differ—one might observe more support for our original hypotheses—in contexts where delegating collection creates more uncertainty about whom to hold accountable or is in fact accompanied by the delegation of more spending authority to local non-state actors. Kananga also differs from many other African contexts in that ethnicity has not been politicized and is generally not a very salient cleavage. In settings where ethnicity is more salient, tax collection by local elites might follow ethnic lines with important implications for accountability.

While these results are surprising and contrary to our original hypotheses, they suggest some important factors for consideration about contexts in which we might expect to see more or less support for the hypotheses. The factor that most likely mitigated support for H1 is that citizens were aware that chiefs remitted revenue to the government and that the government was responsible for spending. We could find more support for H1 in contexts where citizens lack this information or where states actually delegate more revenue and spending authority to local elites along with tax collection. Additionally, one factor that could have mitigated support for H2 is that delegation of tax collection to local chiefs did change their perceived power or capacity relative to the state. We might see more support for H2 in contexts where the chief is perceived to be more independent of the state. This is consistent with the influential argument by Mamdani (1996) that chiefs vary in the extent to which they are incorporated or unincorporated into the colonial government (see also Henn (2020)). In our context, they were incorporated but results could be different where they were not.

The key insight that we document is that delegation of tax collection to local actors does not have adverse effects on accountability, neither in terms of accountability pressures that chiefs are subjected to nor in terms of possible strategic “accountability evasion” by the government via delegation. This runs counter to expectations that the process of tax collection, as well as the collectors themselves, matter for accountability independent of government spending of revenue. It also suggests that local elites, especially chiefs in sub-Saharan Africa, might not be able to evade accountability pressures when tasked with revenue extraction as a large literature has previously argued. Instead, in line with the literature on fiscal decentralization, our results suggest that decentralizing tax collection to lower levels of government might have beneficial consequences for inclusive governance.

Chapter 3

Do Anti-Corruption Policies Attract Votes? Experimental Evidence from India and the U.K.

3.1 Introduction

Anti-corruption programmatic appeals are a common feature of electoral competition across democracies (Bågenholm, 2013*b*; Engler, 2020). In the scholarly literature on the relationship between corruption and voting, it is often argued that activating and politicizing corruption electorally can increase its issue salience and induce voters to vote based on corruption-related considerations, that is to engage in “corruption voting” (De Vries and Solaz, 2017). This activation can happen in numerous ways, through accusations of corrupt practices during election campaigns (Winters and Weitz-Shapiro, 2013), highlighting corruption scandals in the country and increasing media coverage of corruption (Klašnja, Tucker and Deegan-Krause, 2014) or through issue entrepreneurship (Hobolt and De Vries, 2015) by dedicated anti-corruption parties (Bågenholm, 2013*a*) and individual anti-corruption candidates. A successful case of these kinds of strategies is the Aam Admi Party (AAP) in India, the product of a citizen anti-corruption movement, which won a majority of seats in the Delhi State Assembly Elections in 2015 and 2020 promising to fight corruption. Similarly, political candidates such as Volodymyr Zelenskiy and Andrés Manuel López Obrador were elected presidents of Ukraine and Mexico on campaign platforms vowing to address endemic corruption in their countries. Even in low-corruption environments, corruption scandals surface from time to time and can be electorally salient such as in the 2009 UK parliamentary expenses scandal (Eggers, 2014).

However, we have little systematic evidence on whether such anti-corruption policy platforms are in fact electorally successful for candidates and parties, and under which conditions these may override conventional policy items in voters' electoral considerations. Are voter preferences such that voters actually select candidates promising anti-corruption policies in their campaign platform? On the one hand, scholars have argued that good governance should yield electoral returns (Stokes, 1963; Ansolabehere and Snyder, 2000; Green, 2007) and if corruption is such a valence issue, campaign platforms that focus on actively fighting corruption should be electorally successful. Moreover, this form of issue entrepreneurship should also increase concerns about the societal effects of corruption (Klašnja, Tucker and Deegan-Krause, 2014), in turn increasing vote share for parties and candidates campaigning on anti-corruption policy platforms.

On the other hand, there are reasons to doubt that anti-corruption policy promises made during electoral campaigns are electorally successful. The success of anti-corruption parties and candidates running on anti-corruption platforms can either be due to (i) the actual anti-corruption programmatic policies and strategies proposed or (ii) some other element in the anti-corruption platform such as populist rhetoric centred on anti-establishment and anti-corruption sentiments among voters (Busby, Gubler and Hawkins, 2019). Moreover, fighting corruption might simply not generate enough benefits for voters in expectation (Pavão, 2018) and the gains of anti-corruption policies might not be uniformly positive for voters due to clientelism and patronage (Stokes et al., 2013). Moreover, fighting corruption might be perceived as cheap talk, leading voters to condition candidate selection on promises concerning attributable policy outcomes for which effort in office can be evaluated more effectively (Aragonès, Palfrey and Postlewaite, 2007; Harding and Stasavage, 2014; Harding and Goyal, 2019). Finally, whether voters are willing to sacrifice other conventional policy items they care about in order to fight corruption is unclear. Selecting between campaign platforms and policies involves trade-offs with other salient policy issues, and voters may not want to make those trade-offs in favor of fighting corruption, as recent evidence from Mexico on security, violence and corruption suggests (García-Ponce, Zeitzoff and Wantchekon, 2021).

In short, the answer to the question of whether anti-corruption policy promises generate electoral returns is not obvious, and we have little systematic evidence that could provide an answer. Complementing the literature on corruption voting, this

paper asks: what is the effect of anti-corruption policy promises on vote choice?¹ I provide evidence on this question from two conjoint experiments conducted in India and the United Kingdom. The selection of these two cases is not only aimed at probing the scope conditions of the design and results, but also guided by theoretical concerns. That is because the two countries differ markedly in their general level of corruption and citizens’ perceptions of corruption. The comparison of the two samples allows me to test whether anti-corruption policy promises are equally (in)effective electorally in both low-corruption and high-corruption environments. Any theory about the electoral returns to anti-corruption platforms that does not take into account (perceived) levels of corruption is necessarily incomplete.

The conjoint design departs from recently employed conjoint experiments that present respondents with profiles of hypothetical political candidates with “corrupt” or “clean” attributes (Franchino and Zucchini, 2015; Eggers, Vivyan and Wagner, 2018; Breitenstein, 2019; Mares and Visconti, 2020; Agerberg, 2020; Klačnjak, Lupu and Tucker, 2020). Instead, hypothetical candidates campaign on particular policy platforms, including anti-corruption policy, and respondents have to make several trade-offs: between anti-corruption and other conventional policy items such as economic growth and domestic security as well as between those policies and additional candidate characteristics such as gender or party affiliation. These trade-offs reduce the potential for respondents to express support for anti-corruption policies out of social desirability bias or cheap talk (Stephenson, 2015; Incerti, 2020). Furthermore, varying the degree of the effectiveness and associated costs of the proposed policy can make anti-corruption policies attractive to respondents, thus providing stronger evidence of any negative effect that persists under these favorable circumstances (Sniderman, 2018).

Drawing on an array of empirically tested observable implications, I find evidence that runs counter to expectations of the corruption voting literature: the electoral returns to anti-corruption policy platforms are limited. I present four sets of results. First, campaigning on an anti-corruption policy platform, as opposed to economic policy, decreases the likelihood of a candidate being chosen in a hypothetical election. Second, the electoral returns of anti-corruption policies are absent or negative

¹For analytical and empirical clarity, I treat voter support for anti-corruption campaign platforms as one particular form of voting based on corruption considerations. This neglects (i) other forms of corruption voting such as retrospective punishment of corrupt types or prospective selection of non-corrupt types and (ii) other forms of the politicization of corruption during election campaigns such as accusations by opposition parties and judicial reports.

even when (i) fighting corruption entails no further costs and (ii) the particular anti-corruption policy is framed as highly effective in reducing corruption. Third, using outcome data from a hypothetical government budget allocation task post treatment, I provide suggestive evidence that the importance of anti-corruption policies, measured by individual priorities for government spending, is low. Finally, I explore conditions under which anti-corruption platforms receive more or less voter support in high-corruption settings by examining subgroups who are less likely to benefit from current corruption, representing suggestive evidence for effect heterogeneity.

This study makes two contributions to the corruption voting literature. Methodologically, by focusing on the electoral returns of *anti*-corruption policy platforms, I depart from existing conjoint applications featuring “corrupt” or “clean” political candidates. This approach yields conjoint estimates of the electoral returns to anti-corruption policy promises by (i) making voting based on corruption policies attractive via variations in their effectiveness and costs and (ii) reducing the potential for cheap talk by introducing trade-offs with other policies and candidate characteristics. These multiple trade-offs involved increase the confidence in the central finding of this study that the electoral returns of anti-corruption policies are limited both when they are framed as costly and ineffective *as well as when they are framed as non-costly and highly effective*.

The substantive contribution is to present findings that challenge expectations in the existing literature on the theorized positive electoral effects of increasing the salience of corruption during electoral campaigns. Specifically, the evidence suggests that voter preferences over fighting corruption might be less sensitive to political activation efforts and issue entrepreneurship than the corruption voting literature assumes and that not all efforts at making corruption a salient topic during election campaigns are equally likely to resonate with voters. This also illuminates the behavioural microfoundations of claims in the literature that tries to explain the success of new anti-establishment parties, oftentimes campaigning on anti-corruption platforms. The evidence also suggests that the reasons *why* voters do not appear to like anti-corruption policies might not have to do with particular costs associated with fighting corruption or doubts about their effectiveness, but rather with trade-offs with other conventional policy items. Finally, these results suggests that anti-corruption parties and candidates are likely not successful because of the particular anti-corruption *policies* they propose; rather, their anecdotal success could more likely be related to the emotive and symbolic anti-establishment *appeals* that underpin these programmatic platforms.

3.2 Theoretical Framework

Are there electoral returns to campaigning on anti-corruption policy platforms as opposed to other electorally salient policies? I consider competing arguments and hypotheses from a unifying framework of selection and prospective voting as the starting point (Fearon, 1999; Lockerbie, 2008) where three things may matter for voters: benefits from policies, the probability of obtaining those benefits and the salience of the underlying policy issue. The main purpose of this section is to generate, within this prospective voting framework, competing but falsifiable hypotheses about the electoral returns of anti-corruption policy platforms.

3.2.1 The Issue Salience of Corruption in Election Campaigns

Much of the existing literature would expect anti-corruption policy promises to yield electoral benefits for political parties and candidates. In the most general sense, good governance has been shown to affect vote choice, at least in the realm of performance evaluations based on valence issues such as the economy (Duch and Stevenson, 2008; Ashworth, 2012). If fighting corruption lies within the realm of valence politics—where fighting corruption is always good and not fighting it is always bad—this line of reasoning implies that voters should prefer candidates and parties campaigning on such a platform (Stokes, 1963; Ansolabehere and Snyder, 2000; Green, 2007). More specifically, much of the existing literature on corruption voting implies voter preferences that are potentially favorable to anti-corruption platforms. This is theorized to happen when corruption issue entrepreneurship (Hobolt and De Vries, 2015) increases its (sociotropic) issue salience (Klašnja, Tucker and Deegan-Krause, 2014). Of course, the effect might be stronger when those platforms are put forward by dedicated anti-corruption parties (Bågenholm, 2013a) or anti-establishment and independent candidates because their promises to fight corruption might be perceived as more credible (Aragonès, Palfrey and Postlewaite, 2007). However, in general any electoral activation of corruption through a campaign policy platform is expected to increase its issue salience. If voters place more weight on this valence issue, holding constant the credibility of the candidate and their promise, then candidates and parties who campaign on anti-corruption policies are expected to obtain more votes through a process Klašnja, Tucker and Deegan-Krause (2014) call “sociotropic corruption voting”.

These expectations of the existing literature are consistent with studies that have investigated how corruption has been made a salient issue during election campaigns.

Corruption allegations and accusations can solve problems of a lack of information about who is corrupt and who is not (Winters and Weitz-Shapiro, 2013) where access to credible information reduces the vote share of the accused (de Figueiredo, Hidalgo and Kasahara, 2014). In fact, recent studies of new political parties—sometimes with corruption as their main campaign issue—point at newness as a crucial factor for being perceived as a credible actor in terms of fighting corruption (Engler, 2016; Hanley and Sikk, 2016; Haughton, Rybar and Deegan-Krause, 2021), although they do not speak exactly to the question of whether anti-corruption promises are the main reason behind their success. In line with this research, Bågenholm and Charron (2014) show that anti-corruption *rhetoric* is highly successful electorally for opposition parties and new parties, and only in high-corruption contexts.

Activating corruption as a campaign issue is also expected to potentially increase media coverage of the topic, increasing vote shares for those parties and candidates to take an active stance against corruption (Klašnja, Tucker and Deegan-Krause, 2014; De Vries and Solaz, 2017). Finally, anti-corruption measures taken by elite institutions, e.g. the judiciary, just before elections can increase corruption voting by making voters care about integrity (Yair, Sulitzeanu-Kenan and Dotan, 2020).

All this is in line with anecdotal evidence on the success of anti-corruption parties, candidates and movements worldwide. This discussion also implies that when it comes to trade-offs with other electorally salient policy issues, e.g. economic policy, voters could prefer anti-corruption policies. This would also explain why anti-corruption policies would take precedence over pro-growth policies, assuming a zero-sum nature between policy issues. In sum, from this perspective, we would expect voter preferences to be such that they prefer candidates campaigning on anti-corruption measures.

H1: Voters are more likely to prefer a political candidate campaigning on an anti-corruption policy platform over a candidate campaigning on other policy platforms.

3.2.2 Expected Benefits from Fighting Corruption

One problem with the above line of reasoning is that it does not generate consistent predictions as to whether the electoral returns of anti-corruption policies are endogenous to existing levels of corruption. On the one hand, voters in high-corruption environments could have “scandal fatigue” (Klašnja and Tucker, 2013; Klašnja, Little and Tucker, 2016), whereas on the other hand, the issue salience of corruption is much higher in these environments at baseline, potentially leading to more willingness among voters to support anti-corruption measures electorally and to engage in

“sociotropic corruption voting” (Klašnja, Tucker and Deegan-Krause, 2014). Furthermore, due to the scarcity of evidence on the precise question of the electoral returns to anti-corruption policy platforms, we simply do not know whether the apparent anecdotal success of anti-corruption parties, movements and candidates running on anti-corruption platforms is due to (i) the actual anti-corruption programmatic policies and strategies proposed or (ii) some other element in the anti-corruption platform such as populist rhetoric centred on anti-establishment and anti-corruption sentiments among voters (Busby, Gubler and Hawkins, 2019). Finally, it is not clear enough from this line of reasoning whether and why anti-corruption policies might be preferred to other competing and electorally salient policies when trade-offs need to be made. In this section, two alternative explanations are outlined that generate predictions opposite to the one obtained above: a cost-benefit political economy model and a political behavior explanation, both operating within the prospective voting framework.

One reason to expect limited electoral returns to anti-corruption policy platforms is that expected benefits from fighting corruption might be low both in high-corruption and low-corruption equilibria. Under low corruption, the absolute benefits from fighting corruption tend towards zero by diminishing marginal returns as corruption is simply less of a problem, making other policies, such as economic growth policies, more attractive to voters. Economic growth policies are electorally attractive in this scenario as they are needed to (i) attain high growth in low-growth scenarios and (ii) retain high growth in high-growth scenarios to secure employment and wage growth. This is an implication of the economic voting literature that pro-growth policies are electorally rewarded no matter the current state of the economy (Tucker, 2006; Besley, Persson and Sturm, 2010; Ashworth, 2012).

In high-corruption environments, expected benefits from anti-corruption policy platforms might be equally limited. While the pervasiveness of corruption potentially creates high marginal returns to fighting it, actual benefits of anti-corruption policies might not be uniformly positive due to voter benefits from vote buying and patronage (Stokes et al., 2013) or from other economic side effects of corruption (Konstantinidis and Xezonakis, 2013; Fernández-Vázquez and Barberá, 2016; Klašnja, Lupu and Tucker, 2020).²

²Benefits from current corruption of course depend on the precise type of corruption, mainly whether it is grand or petty corruption, and on the type of bribe taker/bribe payer, i.e. bureaucrats, politicians, firms, or all at once. For example, one might argue that corruption in a grab-all-now logic by a dictator with short time horizons (roving bandit) has different effects on citizens benefits than corruption in a grab-over-time logic by stationary bandits which simultaneously at least increases the size of the pie (Fisman and Golden, 2017).

Furthermore, due to pervasive corruption, the uncertainty that any potential benefits from fighting corruption would actually materialize is higher. Above and beyond any baseline uncertainty around policy implementation under high corruption (which would also affect pro-growth policies), there is uncertainty that is specific to anti-corruption policies for at least two reasons. First, fighting corruption is notoriously difficult because of the stickiness of high-corruption equilibria (Fisman and Golden, 2017). Because corruption creates distortions in the policymaking process and is difficult to address with single policies, politicians are believed to be especially incapable of dealing with corruption, more so than with other policy issues (Pavão, 2018). This problem has been compounded by the fact that globalization has made it more difficult to address corruption within national jurisdictions (Gould, 1991; Leiken, 1997; Sung and Chu, 2003). Second, the credibility of the signal of anti-corruption policy platforms might be compromised by the fact that anti-corruption measures in high-corruption environments are often political vendettas against opponents and an effort to signal transparency to international donors (Fisman and Golden, 2017).

By contrast, pro-growth policies potentially yield higher absolute benefits for voters, are not subject to the same policy uncertainty as anti-corruption policies and might therefore be preferred in low-growth and high-growth scenarios alike. This expected benefits framework also suggests that there is likely to be heterogeneity in preferences over policies according to who benefits from current corruption and who does not. The pull of economic growth policies, relative to anti-corruption policies, might be stronger for the poor because pro-growth policies “lift all boats”. Alternatively, opposition supporters benefit less from corruption and clientelism and might therefore be more inclined to prefer candidates campaigning on anti-corruption platforms.

Voter support for anti-corruption policy platforms might also be limited from a political behavior perspective. Various behavioral explanations for limited electoral punishment of corrupt politicians have been put forward in the relevant literature including partisan bias (Anduiza, Gallego and Muñoz, 2013; Xezonakis, Kosmidis and Dahlberg, 2015), in-group loyalty (Solaz, De Vries and de Geus, 2019), institutional features (Xezonakis, Kosmidis and Dahlberg, 2016) or the lack of clean alternatives (Eggers, 2014; Agerberg, 2020). Furthermore, anti-corruption policy proposals might be especially hard to credit on a particular politician or party. Recent arguments suggest that voters are more likely to condition selection of candidates on promises concerning attributable policy outcomes for which effort in office can be evaluated more effectively, such as economic growth, as opposed to reducing corruption (Aragonès,

Palfrey and Postlewaite, 2007; Harding and Stasavage, 2014; Harding and Goyal, 2019).

Regardless of the precise mechanism, both rational choice and political behavior frameworks allow for voter support for anti-corruption policy platforms to be limited, especially vis-à-vis other policy issues such as economic growth, in both high-corruption and low-corruption environments.

H2: Voters are less likely to prefer a political candidate campaigning on an anti-corruption policy platform over a candidate campaigning on other policy platforms.

My research design does not allow me to test exactly the precise mechanism underlying the findings, but observable implications (such as heterogeneity of effects) and scope conditions (such as predictions for high-corruption and low-corruption environments) as predicted by the expected benefits framework can be explored. These observable implications would not only yield further evidence in favor of H2, but also go beyond a simple hypothesis about expected treatment effects. However, the main point is to provide a first set of empirical evidence for a question that has received little systematic research attention so far and where expectations are not obvious and there is disagreement in the literature.

3.3 Experimental Design and Implementation

3.3.1 Setting

To test these competing hypotheses, in April 2019 I conducted two online survey experiments with opt-in subject pools of the Nuffield Centre for Experimental Social Sciences (CESS) in India and the United Kingdom.³ Inferences about the broader voting populations in the UK and India from these convenience samples need therefore be drawn with particular care. For example, the subject pools in both countries are more educated and richer than the respective country average, and the gender ratio in the Indian subject pool is highly biased towards men as can be seen from a comparison of observable covariates presented in Tables B.3 and B.4 in the Appendix.

The selection of India and the United Kingdom serves two purposes. First, it increases confidence in experimental estimates by comparing their direction and magnitude across different contexts. This is especially important with respect to the perceived prevalence of corruption which is higher in India than the UK according to the Corruption Perceptions Index (CPI) which ranks the UK 12 out of 180 countries and India 80/180 in 2019.

³This was two weeks before the Indian Lok Sabha elections.

Additionally, these differences in levels of perceived corruption are the basis for the more substantive and theory-guided rationale for case selection. Systematically testing whether the electoral returns of anti-corruption policies are similar for low-corruption and high-corruption countries (with different levels of economic growth at the time of the experiment) is inferentially difficult and beyond the scope of this paper. However, the selection of India and the UK is designed to partially address this problem. In particular, the UK can be seen to approximate a low-corruption, low-growth world while India tends toward the high-corruption, high-growth equilibrium. In terms of corruption, in India 63% of public service users report to have paid a bribe in the past 12 months according to the Global Corruption Barometer.⁴ India also ranks well above, and the UK well below, global average perceived corruption levels on the Corruption Perceptions Index score. While India is not the perfect representation of a high-corruption equilibrium, it serves as a useful approximation to the limiting case, especially in comparison with the United Kingdom. In terms of growth, annual GDP growth in India was 6.1% in 2018 and 1.3% in the United Kingdom.⁵ Thus, the two countries can be thought of as useful approximations to the low-corruption, low-growth as well as high-corruption, high-growth worlds.

Thus, case selection directly speaks to a prominent debate in the corruption voting literature about differences between high- and low-corruption countries Klašnja and Tucker (2013). Furthermore, case selection was explicitly informed by the concern that anti-corruption rhetoric in low-corruption settings might be less effective and most likely considered cheap talk. Examining results in both high- and low-corruption countries at least descriptively sheds some light on this question.

While a systematic test for similarities across both high-corruption and low-corruption countries is difficult, any similarities with respect to the electoral support for anti-corruption policies across India and the UK would increase the confidence in the hypotheses that in both equilibria, electoral returns to anti-corruption platforms are either large or limited.

3.3.2 Treatments

I employ an experimental conjoint design to test the two competing hypotheses derived above as well as some of their observable implications.⁶ To summarize, the

⁴<https://www.transparency.org/en/countries/india>.

⁵<https://data.worldbank.org/indicator/NY.GDP.MKTP.KD.ZG>.

⁶A pre-analysis plan was informally pre-registered in December 2018 on the author’s github page from which verification and the time stamp is easily retrieved via version control. Deviations from the PAP are outlined in the Appendix. Furthermore, this study has received ethics ap-

observable implications of H2 on the limited voter support for anti-corruption policies are that (i) the effect of anti-corruption policies on candidate selection is negative relative to other policy issues, (ii) the effect is negative or zero even when the anti-corruption policy is framed as highly effective and non-costly, (iii) the importance to voters of fighting corruption is low at baseline compared to other issues, and (iv) the electoral returns to anti-corruption policies vary according to who benefits from corruption.

To measure the potential electoral returns to campaigning on an anti-corruption platform, respondents are exposed to conjoint tables that describe hypothetical candidates who propose policy packages and have certain individual candidate characteristics such as their party affiliation. Table 3.1 is an exemplary conjoint table that could have been realized in the UK experiment, and Table 3.2 depicts all attributes and levels possible for the Indian sample. Political candidates, in the conjoint tasks, differ on policy issues, policy costs, policy effectiveness, as well as candidate credibility, gender, and party affiliation.⁷

Table 3.1: Exemplary Conjoint Table

	Candidate 1	Candidate 2
Policy Issue	Anti-Corruption	Economy
Policy Costs	No costs	Raising income taxes
Policy Effectiveness	Has been shown to perform well	Has been shown to have little effect
Candidate Credibility	Has a reputation for keeping campaign promises	Has a reputation for not keeping campaign promises
Candidate Gender	Male	Female
Candidate Party	Labour	Conservative

The main attribute of interest is policy issue which can take on the levels {anti-corruption, economy, domestic security} introducing trade-offs with other salient policy issues.⁸ To make these policy issue treatments comparable across contexts and to avoid priming respondents on particular aspects of a certain exactly specified policy, the policies proposed by the hypothetical candidates were framed in a more general

proval from the University of Oxford Departmental Research Ethics Committee (DREC), Ref. No.: SSH.DPIR.C1A.15.066.

⁷Note that this diverges from recent conjoint experimental approaches that let subjects choose between candidates who are either accused to be corrupt or shown not to be corrupt.

⁸Although not theorized, domestic security/fighting crime serves as an empirical benchmark against which to anchor any effects of an anti-corruption campaign platform. As Fisman and Golden (2017) note, crime is one of the most salient issues in public opinion worldwide. Furthermore, García-Ponce, Zeitzoff and Wantchekon (2021) have recently shown that voters in Mexico are willing to accept corruption for a reduction in violence.

way. The implicit theoretical and empirical assumption here is that policy issues are zero sum: no candidate can ever propose a policy that is pro-growth *and* anti-corruption. This assumption might be violated if no candidate can ever ignore pro-growth policies in their electoral campaigns. Furthermore, if growth can be achieved through various different mechanisms, one of which might be fighting corruption, then the design might be underestimating the impact of anti-corruption policies and put limitations on the conclusions about the electoral costs of anti-corruption platforms. Future work should build on this limitation by, for example, combining different policy packages in the conjoint attribute levels. It is important to note, however, that real-world trade-offs in campaign designs are made regularly, especially when it comes to policy emphasis. In addition, budget constraints often force political candidates to take a stance on one policy issue and not another. This is confirmed in the budget allocation analysis where voters are willing to make budgetary trade-offs between fighting corruption and promoting economic growth.

In addition, the attributes policy costs and policy effectiveness aim to test whether respondents are willing to incur costs or tolerate ineffective policies for their preferred policy issue, reducing the potential for social desirability bias (Incerti, 2020). Conversely, the effectiveness, costs and credibility attributes control voter concerns about expected benefits and costs of anti-corruption policies, and thus can make it attractive for voters to vote based on corruption considerations.⁹ The component interaction effects between the policy issue and its costs and effectiveness can uncover voter willingness to select anti-corruption platforms under favorable circumstances.¹⁰ For example, persistent negative effects even under these favorable circumstances increase the confidence in H2 about the limited electoral support for anti-corruption platforms, following the dosage logic as explicated in Sniderman (2018). Candidate partisanship, gender and credibility are added to reduce the potential for respondents to project

⁹One design limitation is that we have no way of knowing whether subjects in the high-corruption setting are actually convinced by any positive effectiveness or credibility framing of a particular policy, given high baseline distrust in political institutions. Furthermore, specific anti-corruption parties such as the Aam Admi Party (AAP) in India are not included in the party attribute, which might affect the overall credibility of a proposed anti-corruption policy. However, it was important to include party labels to make the conjoint task more realistic. Furthermore, it remains unclear whether opposition and challenger parties might be considered more credible without the possibility of establishing a reputation while out of government.

¹⁰One possible concern is that policy costs are not independent of the policy issue. For example, the disclosure of private data is much more relevant for improving domestic security than for stimulating economic growth. However, the benefits of completely independent randomization without attribute constraints outweigh these concerns.

attributes that are salient to them but missing in the conjoint onto attributes in the conjoint design (“aliasing”).

Table 3.2: India Conjoint: Attribute Levels

Attribute	Level	Text in table
Costs	Public money/taxes	Raising income taxes
	Opportunity costs	Less money for other issues
	Personal data	Disclosure of private data
	No costs	No costs
Effective	Effective	Has been shown by experts to perform well
	Not effective	Has been shown by experts to have little effect
Issue	Anti-corruption	Anti-corruption
	Economy	Economy
	Domestic Security	Domestic Security
Credibility	Credible	Has a reputation for keeping campaign promises
	Not credible	Has a reputation for not keeping campaign promises
Gender	Female	Female
	Male	Male
Party	Incumbent party	Bharatiya Janata Party
	Main opposition party	Indian National Congress

To measure the potential electoral returns to campaigning on an anti-corruption platform, a conjoint design is preferable to a standard vignette design. The trade-off nature introduced by conjoint tasks is ideal for measuring the effect of different policy proposals and campaign promises that also differ on other attributes. Therefore, the potential for social desirability bias and non-costly support for fighting corruption is reduced by (i) opportunity costs of not choosing a candidate campaigning on other, perhaps more important, policy issues such as the economy, (ii) costs associated with each policy, and (iii) the effectiveness of each policy proposal.¹¹

3.3.3 Outcome Measures

Candidate selection is measured in a forced-choice design as suggested by Hainmueller, Hopkins and Yamamoto (2014). Respondents are asked, after having read each conjoint table, “If you had to choose between them, which of these candidates would you vote for in the upcoming election?”¹² Choice-based conjoint analysis attempts to

¹¹Note that hypothetical bias is reduced, but not eliminated by measuring a willingness to pay for policy proposals. Costs in the conjoint setup are still purely hypothetical (Incerti, 2020).

¹²The survey also asked respondents to rate both candidates on a scale from 1 to 7. However, due to technical difficulties with transposing data matrices into a conjoint-friendly format, I did not analyze responses to these rating questions here.

approximate real-world voter decision making and captures stated voter preferences over candidates (Hainmueller, Hopkins and Yamamoto, 2014; Hainmueller, Hangartner and Yamamoto, 2015).¹³

Among the set of prognostic pre-treatment covariates included such as respondents’ age, gender, education and ideology, respondents also report their annual household income as well as their party support conditional on party identification. This allows to tentatively explore treatment effect heterogeneity by income and incumbent party support as theorized above and pre-registered as heterogeneous effects.¹⁴ Post treatment, respondents are asked to plan a hypothetical budget for the upcoming fiscal year and allocate resources that should be spent on each of six issues. This is to test an additional observable implication of the two competing hypotheses to bring data from yet a different choice task on the problem. The full surveys including exact question wordings for both countries are reproduced in the Appendix.

3.3.4 Analysis

The conjoint experiment draws on $N = 212$ respondents in India and $N = 218$ in the UK who evaluate in each of $K = 6$ choice tasks $J = 2$ hypothetical candidate profiles differing across $L = 6$ attributes. Because conjoint analysis pools results from all 6 choice tasks, it draws on subject-performed $212 \times 6 = 1,272$ choice tasks with 2,544 profiles evaluated in India and $218 \times 6 = 1,308$ choice tasks and 2,616 profiles evaluated in the UK. Conjoint profiles were randomly generated by completely independent randomization, drawing attribute levels from a uniform distribution without randomization weights for any given attribute. That is, the realization of some attribute levels is not constrained by other attribute levels.¹⁵

A range of diagnostic tests in the Appendix presents evidence in favor of no profile-order effects or carryover effects and also shows equal display frequency of attribute levels in the two experiments. Due to attribute order constraints which fix

¹³Note that the claim of this paper is not about the attribute composition of the likely *winner* of an election or the representativeness of the attributes and their levels of the larger voting populations in India and the UK. As Abramson, Koçak and Magazinnik (2019) and de la Cuesta et al. (2020) show, the impossibility of these two interpretations of conjoint outcomes limits the external validity of conjoint designs.

¹⁴Respondents were also asked whether they had received any material benefits in exchange for voting for a particular candidate in the last election. Due to concerns about possible response bias, I do not analyze this question further.

¹⁵Note that, due to the lack of randomization constraints, it was possible that two candidates from the same political party occurred in the same choice task. Even though this diminishes the external validity of the design, this is a minor concern given the returns from completely independent randomization.

all policy and candidate attributes to always be in the same order, a test for primacy effects is impossible. Yet, those primacy effects are unlikely given the limited number of attributes, allowing for sufficient complexity to avoid aliasing of attributes while reducing the likelihood of overwhelming respondents (Bansak et al., 2019).

There was no attrition on the outcome variable. I employed a treatment-irrelevant factual manipulation check (FMC-TI) following Kane and Barabas (2019) that asked respondents after completion of all six tasks about material that was not manipulated across treatments, in this case the number of attributes per profile. In India, 57.1% answered this question correctly, while in the UK, 66.5% had the correct answer. The question wording of the FMC seemed to have been slightly confusing for subjects, and therefore, as well as to avoid post-treatment bias (Acharya, Blackwell and Sen, 2016), subjects giving incorrect answers were not dropped from the analysis.

Three estimands are of particular interest. First, the *average marginal component effect* (AMCE) of the levels of the policy issue attribute on candidate choice identifies the effect of campaigning on an anti-corruption platform vis-à-vis other policy issues such as the economy or domestic security. The AMCE quantifies how different levels of the l th attribute, for example variations in policy issue, of profile j influence the probability that the profile is chosen. We obtain the marginal effect of attribute l averaged over the joint uniform distribution of all possible attribute combinations. Thus, this estimand quantifies the change in the probability that a profile would be chosen if the level of its l th component were changed from the baseline level t_0 (control group) to the level of interest t_1 (treatment group), in this case to the anti-corruption policy platform.

Second, to test additional observable implications of the two competing hypotheses, I am interested in the *average component interaction effect* (ACIE) which quantifies the size of interactions when the causal effect of one attribute (e.g. policy issue) varies depending on what value another attribute (e.g. policy costs) is held at: what is the effect of a particular policy issue on candidate choice when we consider, for example, costly policies in the form of tax increases? From this estimand, we identify the percentage point difference in AMCEs of a particular policy issue such as an anti-corruption platform between different forms of costs associated with this policy.

Finally, to explore the observable implication of the rational choice approach (H2) that voter support for anti-corruption policies varies by benefit status from the corruption status quo in high-corruption contexts, an important estimand is the *conditional AMCE*. It identifies the effect of an attribute interacted with a respondent's

background characteristic. This corresponds to modelling heterogeneous treatment effects.

Following Hainmueller, Hopkins and Yamamoto (2014), the AMCE is estimated from conjoint data by regressing the choice made by respondent i in their k th choice task about profile j , Y_{ijk} , on a set of 6 categorical variables (1 per attribute) indicating the levels of each attribute with the simple regression estimator $Y_{ijk} = 1 + W_{ijkl}$, where 1 is an intercept and W_{ijkl} a vector of $D_l - 1$ dummy variables for the levels of T_{ijkl} except the one for the baseline level t_0 . Empirically, this amounts to simply regressing, in a logit model, the binary outcome variable candidate choice on the set of factor variables that represent the attribute levels and clustering the standard errors by respondent as the choice tasks are nested within subjects. Confidence intervals are based on these standard errors. We can extend the estimating equation to the cases of ACIE and conditional AMCE by further sub-classification of the sample into strata defined by either the attribute with which the interaction is of interest (ACIE) or covariates X_i (conditional AMCE).

In the Appendix, all results are replicated using marginal means (MMs) and its subgroup and interaction variants for additional robustness checks. The marginal mean of a conjoint attribute level is the mean outcome across all appearances of a particular conjoint attribute level averaged over all other attributes. In forced-choice designs with two profiles such as the one employed here, MMs by design are on average 0.5 where values above 0.5 indicate increases in profile favorability and values below 0.5 indicate decreasing profile favorability (Leeper, Hobolt and Tilley, 2019). Throughout, both AMCE and marginal means analysis are considered.

3.4 Results

In this section, I present four sets of results that corroborate H2 of limited electoral returns of anti-corruption policies: (i) the effect of anti-corruption policies on candidate selection; (ii) the interaction effects of policy costs and effectiveness; (iii) the baseline importance of anti-corruption measures in a budget allocation task; (iv) the effects of anti-corruption policies conditional on respondent income and party support.

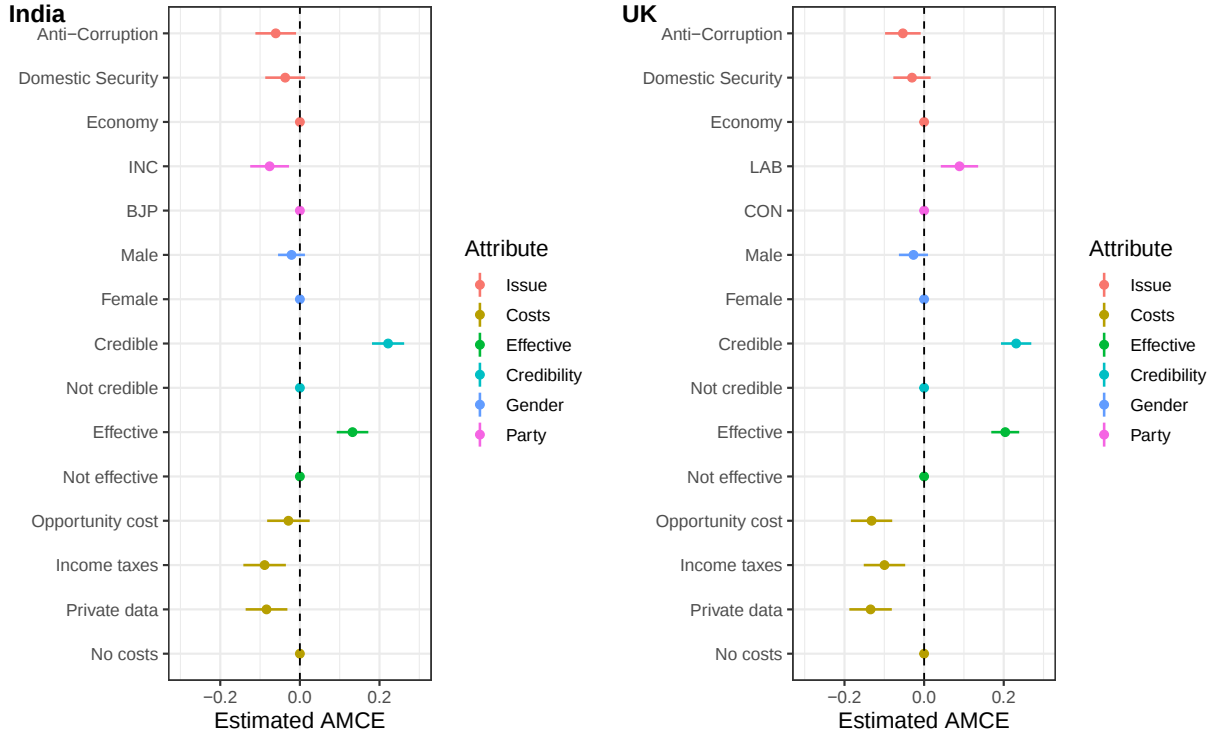
3.4.1 Anti-Corruption Campaign Platforms

The main estimand of interest is the average marginal component effect of an anti-corruption policy platform on the probability of being chosen as a political candidate.

Recall that the AMCE represents the change in the probability that a particular profile j in choice task k is chosen over the other profile when moving from t_0 (baseline or comparison attribute level) to t_1 , i.e. the attribute level of interest. Figure 3.1 reports the AMCEs from a binary logit model alongside 95% confidence intervals based on clustered standard errors. In both samples, the AMCE for the anti-corruption policy attribute is negative and significantly different from zero at $p < 0.05$. That is, campaigning on an anti-corruption platform (t_1) reduces the probability of being chosen as a candidate by 6.1% (India) and 5.4% (UK), compared to the counterfactual where the same candidate with the same characteristics had proposed an economic policy to stimulate growth (t_0). For comparison, campaigning on a domestic security platform does not invoke such statistically significant effects on candidate choice. These effects are remarkably similar across both low-corruption, low-growth as well as high-corruption, high-growth contexts.

It is worth noting that these negative effects of anti-corruption platforms are present even in the face of other strong predictors of candidate choice, such as the credibility of the candidate and the effectiveness of the policy. These two are the strongest effects in substantive terms, also reaching statistical significance at $p < 0.05$. First, proposing an effective policy that has been shown to work well to address the underlying policy issue, as compared to a non-effective policy, increases the probability of being chosen as a candidate by 13.3% (India) and 20.3% (UK). Second, having a reputation for keeping campaign promises instead of having a non-credible candidate reputation increases the probability of being chosen as a candidate by 22.2% (India) and 23.1% (UK). The reverse is true for the attribute policy cost. Proposing a policy that requires an increase in income taxes or the disclosure of citizens' private data decreases the probability of being chosen as a candidate when compared to the counterfactual of no costs by around 10% in both samples. The marginal means analysis (Figure B.6 in the Appendix) confirms these conclusions, with the exception that the marginal mean for anti-corruption in the UK is not significantly different from 0.5 at $p < 0.05$. Finally, it is worth noting that these negative effects of anti-corruption platforms are present in a sample so strongly in favor of the ruling BJP party which could presumably be the target of anti-corruption efforts; the AMCE on the main opposition party INC is significantly negative. Overall, this represents evidence in favor of H2 which predicted limited voter support for anti-corruption policy promises.

Figure 3.1: AMCEs of the Candidate Choice Experiment



Notes: This figure displays estimated Average Marginal Component Effects (AMCEs) for the candidate choice experiment in India (left panel) and the U.K. (right panel). These are results from binary logit models where the outcome, candidate choice, is regressed on series of categorical variables indicating policy issues, costs, and effectiveness as well as candidate credibility, gender, and party. 95% confidence intervals are based on standard errors clustered at the respondent level. All analyses are implemented with the R package `cregg`.

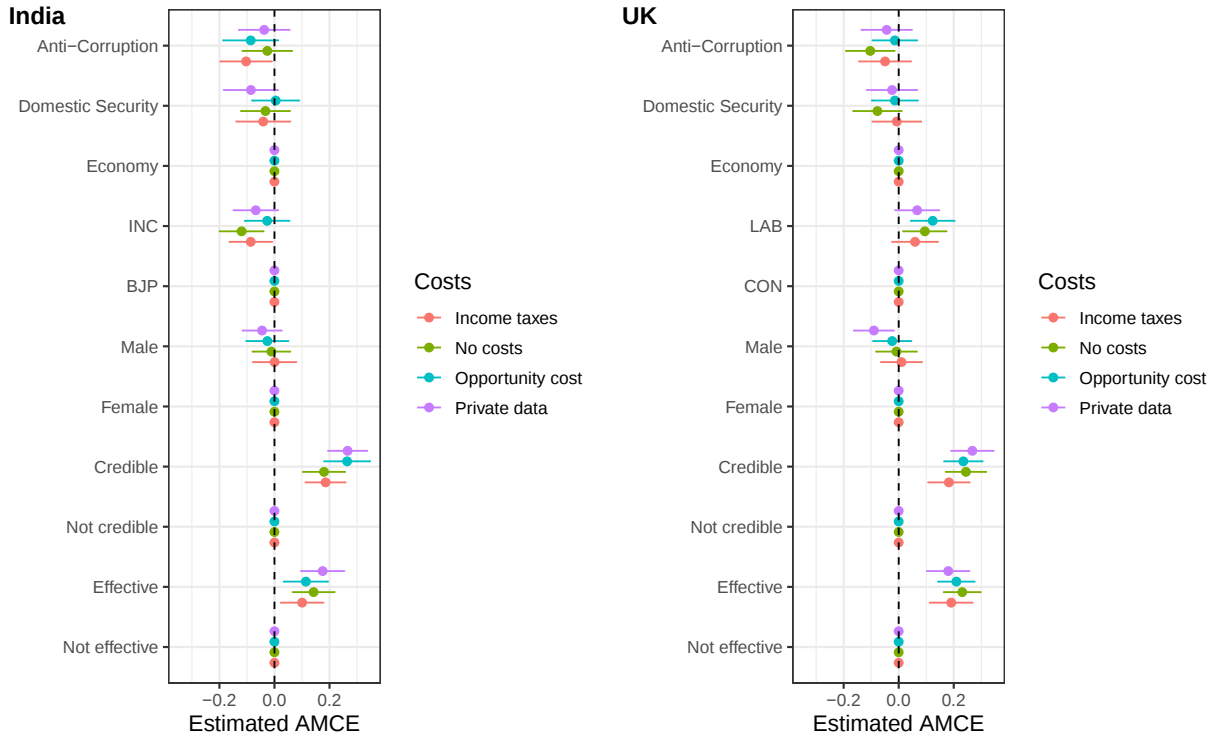
3.4.2 Anti-Corruption Platforms, Costs and Effectiveness

One observable implication of H2 (limited voter support for anti-corruption policies) is that the negative effect of anti-corruption policies persists, or at least does not turn into a significant positive effect, *even if those policies are non-costly or highly effective*. To bring data on this observable implication, Figure 3.2 and Figure 3.3 present average component interaction effects. Recall that the ACIE for a given attribute level represents the average % difference in the AMCEs of component l , in this case policy issue, when holding constant different levels of some other component, in this case policy costs or policy effectiveness.

Consider first the effect of anti-corruption platforms depending on the costs associated with such a policy. Even if the proposed anti-corruption policy entails no costs, campaigning on such a platform as opposed to an economic one decreases the proba-

bility of being chosen as a candidate by 10.4% at $p < 0.05$ (UK) and has no effect at all on the probability of candidate choice in India, where the ACIE is close to zero and not statistically significant. This suggests that promising to fight corruption is not an electorally successful strategy even if voters do not have to pay for it. Conversely, the lack of electoral returns to anti-corruption platforms is especially pronounced when voters would have to pay for fighting corruption by paying higher income taxes. The effect of costly (that is, paid for by higher income taxes) anti-corruption policies is negative and statistically significant in India. The analysis of marginal means for the issue/costs interaction in Figure B.7 in the Appendix yields the same conclusions.

Figure 3.2: ACIEs for Varying Policy Costs



Notes: This figure displays estimated Average Component Interaction Effects (ACIEs) for the candidate choice experiment in India (left panel) and the U.K. (right panel) for varying policy costs. These are results from binary logit models where the sample is subset to strata defined by the attribute policy costs and the outcome, candidate choice, is regressed on series of categorical variables indicating policy issues and effectiveness as well as candidate credibility, gender, and party. 95% confidence intervals are based on standard errors clustered at the respondent level. All analyses are implemented with the R package `cregg`.

The same is true for the effect of anti-corruption platforms depending on the effectiveness of the policy. Even if the proposed anti-corruption policy is framed as a

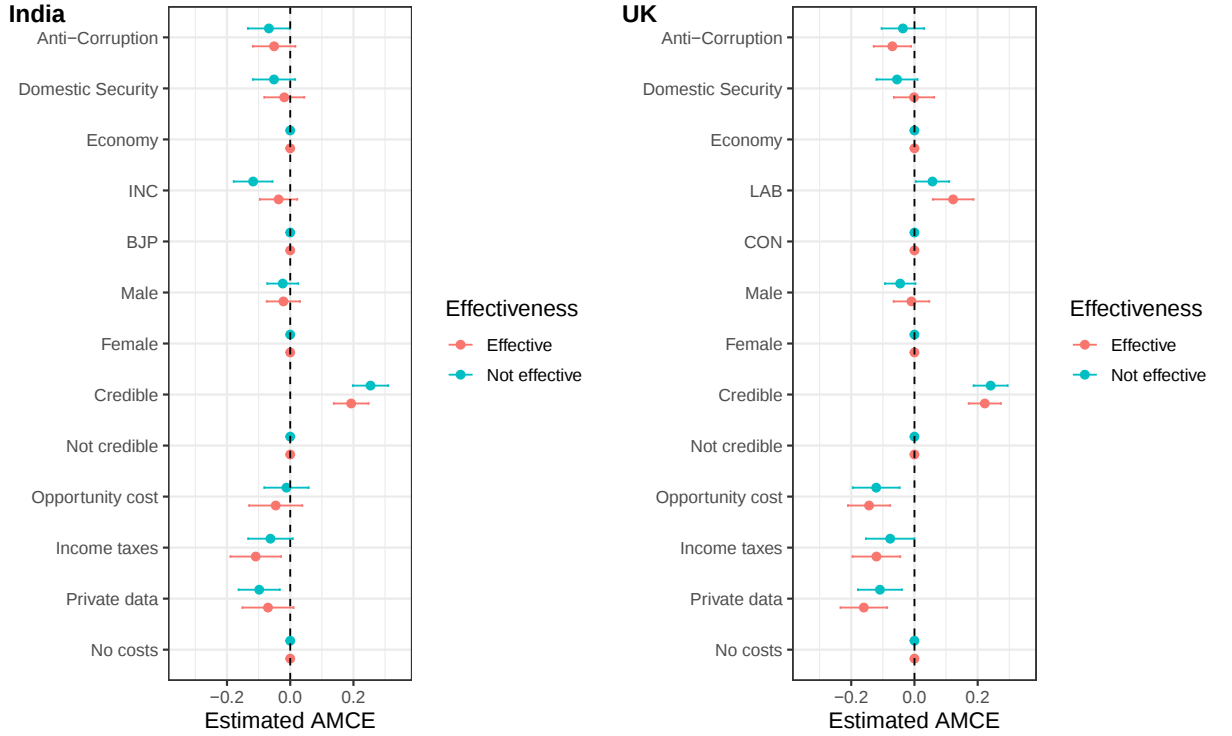
highly effective measure to address corruption, campaigning on such a platform as opposed to an economic one decreases the probability of being chosen as a candidate by 6.9% at $p < 0.05$ (UK) and has no statistically significant effect on the probability of candidate choice in India where the 95% confidence interval includes zero. Conversely, in India non-effective anti-corruption campaigns are statistically significantly less likely to be selected than non-effective economic policies. Taken together, these results imply that even when trying to address possible voter concerns about the effectiveness or costs of fighting corruption by framing it respectively, anti-corruption platforms do not yield positive electoral returns. The same conclusions are reached by the analysis of marginal means of the issue/effectiveness interaction reported in Figure B.8 in the Appendix, further supporting H2 of limited voter support for anti-corruption policies.

3.4.3 Budgetary Salience of Anti-Corruption Policies

Results from yet a different strategy for measuring the salience of corruption increase our confidence in the main results reported above, corroborating the hypothesis that voter support for anti-corruption campaign platforms is limited. The inclusion of anti-corruption policy platforms as one possible policy issue in the conjoint experiment did not seem to have increased the salience of fighting corruption in other kinds of choice tasks, such as the hypothetical budget allocation task employed post treatment.

In both subject samples, fighting corruption does not appear to be a top policy priority according to the share of a hypothetical government budget allocated to each of six policy issues. Figure 3.4 ranks those policy issues according to the median % of budget allocated to a given issue and depicts the entire distribution and their respective mass points in a jitter plot. Corruption, post treatment, is the second least important issue to address in the Indian sample, and the least important in the UK sample, with 15.1% (India) and 11.3% (UK) mean budget share. It also confirms the appropriateness of the two “competitor issues” economy and domestic security as additional policy attribute levels in the conjoint experiments. The economy is in both samples the second most important issue (19.1% in India and 18.7% in the UK) which increases confidence in the experimental findings on the AMCEs and hypothesis 2 that campaigning on a platform to fight corruption attracts less votes than campaigning to address other salient issues. However, these results do not involve any experimental manipulation which cautions against a strong interpretation.

Figure 3.3: ACIEs for Varying Policy Effectiveness



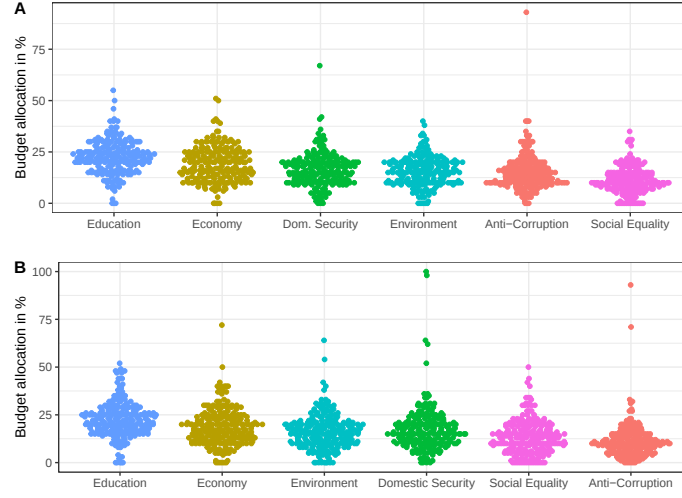
Notes: This figure displays estimated Average Component Interaction Effects (ACIEs) for the candidate choice experiment in India (left panel) and the U.K. (right panel) for varying policy effectiveness. These are results from binary logit models where the sample is subset to strata defined by the attribute policy effectiveness and the outcome, candidate choice, is regressed on series of categorical variables indicating policy issues and costs as well as candidate credibility, gender, and party. 95% confidence intervals are based on standard errors clustered at the respondent level. All analyses are implemented with the R package `cregg`.

3.4.4 Treatment Effect Heterogeneity

Additionally, the political economy rationale outlined above suggests that, in high-corruption equilibria, there is likely variation in voter support for anti-corruption policies in terms of beneficiaries and non-beneficiaries from the status quo. In particular, those who would presumably have the highest marginal returns from anti-corruption policies such as the poor might, perhaps counterintuitively, shy away from those policies because potential economic gains from anti-corruption policies are seen as highly uncertain.

A first way to tentatively explore this possibility is to examine heterogeneity in treatment effects by respondent income and partisanship both of which were pre-

Figure 3.4: Percentage Allocated for a Hypothetical Government Budget



Notes: This figure displays the percentages allocated for a given issue in a hypothetical government budget allocation task. Points are distributed within a kernel density estimate of the distribution with offset determined by quasirandom Van de Corput noise. They are ranked according to the median percentage allocation. Panel A is the Indian subject pool, Panel B the UK subject pool. Respondents were asked the following: “Imagine you are planning the federal budget for the upcoming budgetary year. Please allocate percentages of money that should be spent on each issue. The percentages must add up to 100%.”

registered as potentially moderating treatment effects.¹⁶ The pull towards policies with more certain benefits might be stronger for the poor and non-incumbent partisans as they profit from lifting all boats through economic growth more than beneficiaries from the current status quo such as the rich or incumbent supporters. The expectations are that (i) the negative effect of anti-corruption platforms on candidate selection is stronger among the poor and among subjects who do not identify with the incumbent party and (ii) this effect heterogeneity is absent in the UK, a low-corruption country.

Subjects’ reported household income category in India was recoded as low for subjects with household income between less than 60,000 and 89,999 Indian rupees, as middle for those with income between 90,000 and 239,999 Indian rupees and high for subjects with income higher than 240,000 Indian rupees. Similarly, for the UK,

¹⁶This is by no means a systematic and comprehensive test of the observable implication that those who benefit less from the current corruption status quo are still less likely to support anti-corruption platforms. Heterogeneous treatment effects are, by design, non-experimental in nature as income groups as well as party identifiers are formed before random realization of profiles. Additionally, the limited sample size and partially inconclusive results as reported below caution against an overly confident interpretation of this exploration into treatment effect heterogeneity.

low-income subjects are those with income less than 10,000 and 19,999 British pounds, middle-income subjects are those with incomes between 20,000 and 59,999 and high-income subjects are those with income higher than 60,000 British pounds.¹⁷

In line with theoretical expectations outlined above, the AMCE on anti-corruption is negative and statistically significant for low-income subjects in India as shown in Figure 3.5. Note here that the baseline policy issue is now domestic security to directly compare effects for anti-corruption and economic policies. Among low-income subjects, campaigning on an anti-corruption policy as opposed to domestic security decreases the probability of being chosen as a candidate by 14.8%. This difference in AMCEs between only the high and low income subjects is statistically significant with $p < 0.05$ (and the 95% confidence interval does not include zero) as can be seen from Figure B.4 in the Appendix which depicts the differences in AMCEs across attributes for high and low-income group subjects. Note that the point estimate for the marginal mean for anti-corruption policies among low-income subjects in the India sample is lower than 0.5, but this difference is not statistically significant as reported in Figure B.9 in the Appendix. This income gradient in voter support for anti-corruption policy platforms is consistent with recent findings by Klačnja, Lupu and Tucker (2020) that voters with lower wealth sanction corrupt politicians less than high-wealth voters.

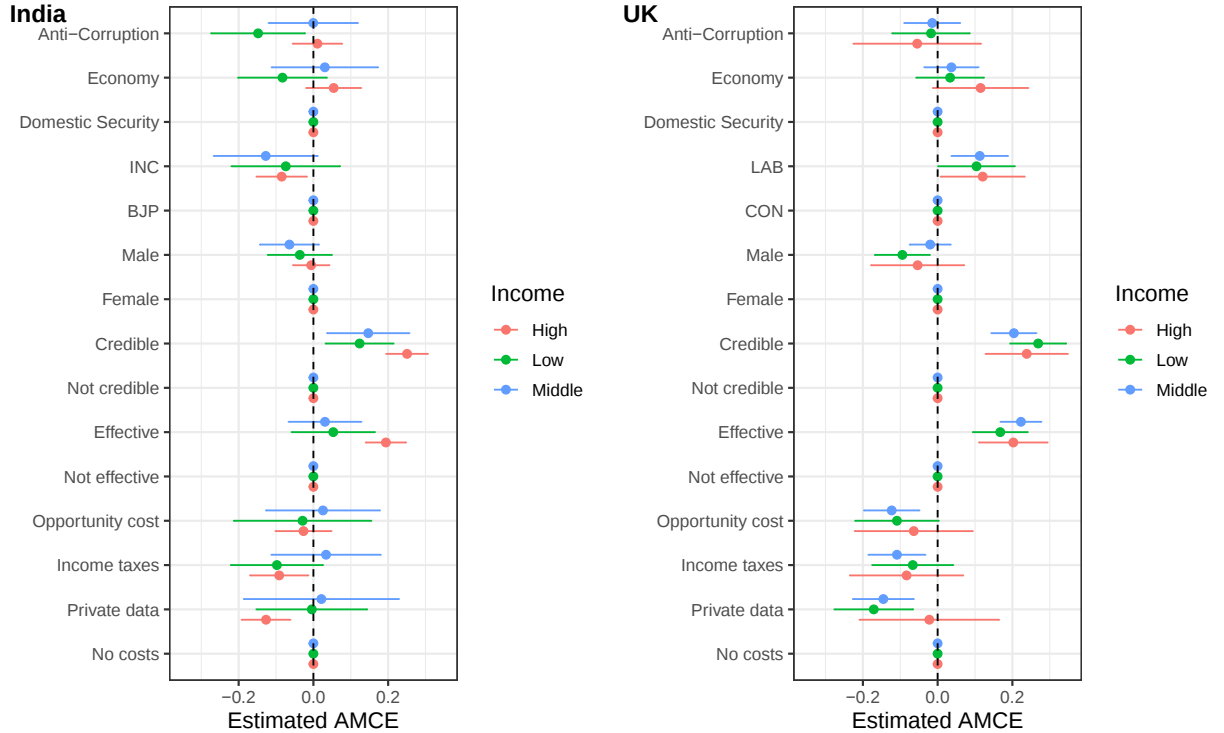
Furthermore, this income gradient in anti-corruption policy preferences is not present in the UK sample. Low-income subjects are not less likely to select a candidate campaigning on an anti-corruption policy than an otherwise similar candidate campaigning to stimulate growth, and there is no statistically significant difference in AMCEs for the anti-corruption attribute between high and low income subjects. This conclusion about the UK sample is corroborated by the marginal means analysis by income as reported in Figure B.9 in the Appendix.¹⁸

The second source of possible variation in effect heterogeneity is partisanship of respondents, specifically whether they self-identify with the incumbent party. However, the results can be interpreted as suggestive at best due to limited sample size

¹⁷The results are robust to a discretization involving no middle income category as reported in Figure B.11 in the Appendix.

¹⁸However, additional caution about the confidence in these results is in order above and beyond limited sample size for subgroup analysis. The non-experimental nature of covariate conditioning of treatment effects potentially explains some portion of the heterogeneity of AMCEs by respondent income for two reasons. A first possibility is that income might just be correlated with education. The poor might be less educated on average, and that could explain their preference against fighting corruption. A second explanation is the degree of informality of the two economies and the ease of access to public services without paying a bribe. Both these possibilities call for more systematic investigation which is beyond the scope of this paper.

Figure 3.5: AMCEs Conditional on Household Income



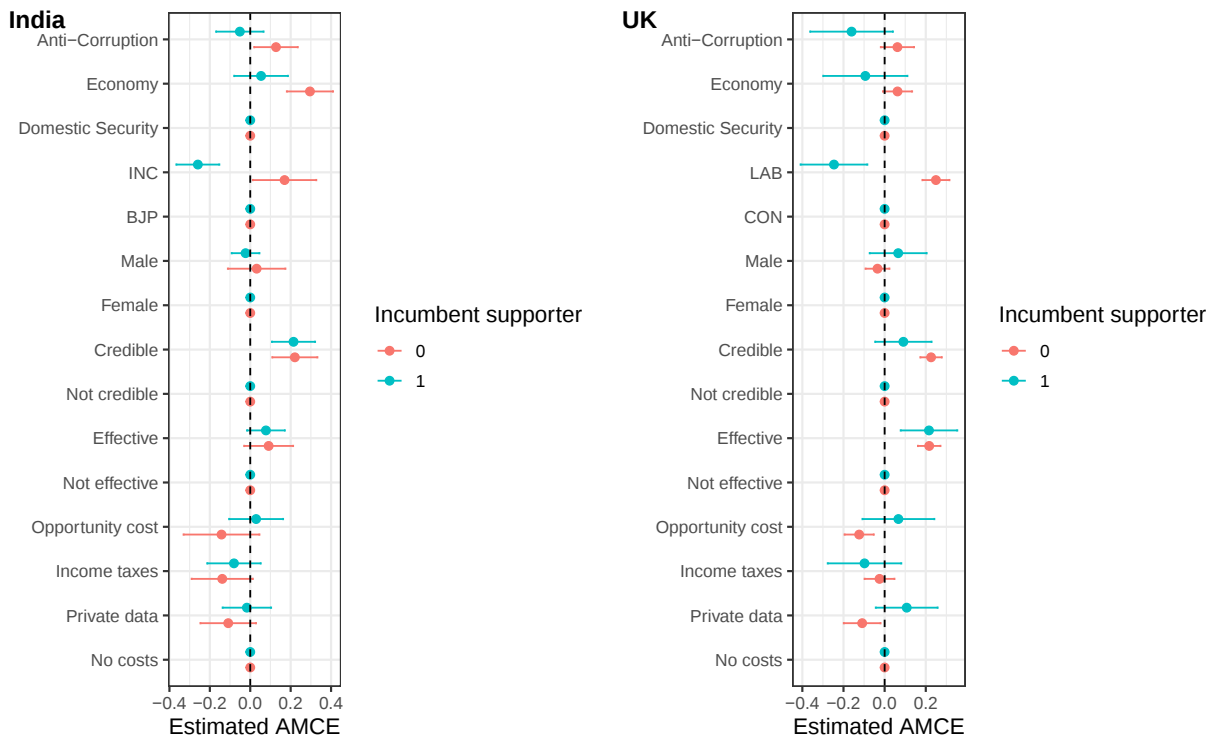
Notes: This figure displays estimated Average Marginal Component Effects (AMCEs) for the candidate choice experiment in India (left panel) and the U.K. (right panel) conditional on household income. These are results from binary logit models where the sample is subset to strata defined by household income and the outcome, candidate choice, is regressed on series of categorical variables indicating policy issues, costs and effectiveness as well as candidate credibility, gender, and party. 95% confidence intervals are based on standard errors clustered at the respondent level. All analyses are implemented with the R package `cregg`.

among the subset of respondents who identify with any political party. In India, 36 respondents identified as supporters of the BJP, and 21 were coded as supporting a different party. In the UK, 16 respondents identified as supporters of the Conservative Party, and 73 were coded as other party supporters. Nevertheless, we observe heterogeneity of marginal component effects conditional on whether subjects identify with the incumbent party or not, although at odds with the expectations formulated above.

As Figure 3.6 shows, the effect of an anti-corruption platform vis-à-vis a domestic security platform on candidate choice is *positive* and statistically significant among *non*-incumbent party supporters in India, running counter to the expectation that non-incumbent respondents might shy away from supporting anti-corruption policies

due to a stronger pull by more certain benefits from economic growth policies. By contrast, these empirical results suggest that those who are less likely to benefit from corruption through vote buying or clientelistic exchanges are more likely to support the fight against corruption electorally. This heterogeneity in AMCEs is absent in the UK where none of the subgroups is more or less likely to support an anti-corruption platform over a domestic security platform. Furthermore, the difference in AMCEs in India is statistically significantly different from zero, as Figure B.5 in the Appendix shows. This is true for the UK too, although none of the two conditional AMCEs are significantly different from zero.

Figure 3.6: AMCEs Conditional on Incumbent Party Support



Notes: This figure displays estimated Average Marginal Component Effects (AMCEs) for the candidate choice experiment in India (left panel) and the U.K. (right panel) conditional on incumbent party support. These are results from binary logit models where the sample is subset to strata defined by incumbent party support and the outcome, candidate choice, is regressed on series of categorical variables indicating policy issues, costs and effectiveness as well as candidate credibility, gender, and party. 95% confidence intervals are based on standard errors clustered at the respondent level. All analyses are implemented with the R package `cregg`.

However, note that the subgroup marginal means analysis conditional on identification with the incumbent party as reported in Figure B.10 for both samples does

not support this claim; the two marginal means for the anti-corruption platform are not significantly different for incumbent supporters and non-incumbent respondents. Overall, the limitations discussed in this section — limited sample size in subgroups, non-experimental nature of covariate conditioning, and partially inconclusive results with respect to marginal means — only reinforce the need for further and more systematic analysis into variation of corruption voting by status as beneficiary from corruption.

3.5 Discussion

This paper has shed light on a form of electoral activation of corruption that has so far received little systematic research attention: anti-corruption policy platforms. Testing competing hypotheses about the electoral returns to anti-corruption policy promises during campaigns, I have shown empirically with a wide array of observable implications that the electoral returns to anti-corruption policy promises are limited at best when compared to campaign platforms based on other policy issues. Multiple trade-offs involved increase the confidence in the central finding of this study that the electoral returns of anti-corruption policies are limited both when they are framed as costly and ineffective *as well as when they are framed as non-costly and highly effective*.

I find that voter preferences over fighting corruption might be less elastic to political activation efforts and issue entrepreneurship than the corruption voting literature assumes. The evidence also suggests that *why* voters do not appear to like anti-corruption policies might not have to do with particular costs associated with fighting corruption or doubts about their effectiveness, but rather with trade-offs with other conventional policy items. I have also tentatively explored conditions under which voter support for anti-corruption platforms in high-corruption environments is less likely. The theoretical argument and exploratory evidence on treatment effect heterogeneity suggest that electoral support likely varies with the extent to which subgroups of voters benefit from the corruption status quo, although these results are not entirely conclusive and partially inconsistent with theoretical expectations. Given the clear empirical limitations of these particular results, more systematic research is needed to examine this heterogeneity.

Examining electoral support for anti-corruption platforms also yields new insights about our understanding of the relationship between corruption and voting more

broadly. The central result of this study is that even under most permissive circumstances, voter support for candidates promising to fight corruption is lower than for candidates who campaign on, for example, economic policy. This suggests that the low incidence of corruption voting might be more common to a wider set of political contexts and circumstances and is not only found in the (lack of) retrospective punishment of corrupt politicians which much of the literature has focused on (Fisman and Golden, 2017; De Vries and Solaz, 2017; Dunning et al., 2019; Incerti, 2020).

However, this study has a broader set of limitations, both substantively and methodologically. As for the former, the results only speak to the demand side of the politicization hypothesis, not to the supply side on how politicians respond to those electoral incentives, as the conjoint design does not allow for a thorough estimation of the sequential nature of the campaign game. Additionally, I am not testing arguments about electoral accountability as such, nor about the effect of anti-corruption platforms on retrospective punishment or prospective selection of corrupt candidates; the goal here is merely to show that even under favorable circumstances, voter support for anti-corruption platforms is largely limited. Most importantly, this paper does not directly test which of the mechanisms that led to opposite predictions about the electoral returns to anti-corruption platforms is at work. There is suggestive evidence for a political economy explanation as (i) results are similar for both high-corruption and low-corruption contexts and (ii) there is some evidence of effect heterogeneity according to benefit status from corruption. However, this set of results has clear methodological limitations that were outlined above and future research is needed to pinpoint exactly which mechanism is leading to reduced voter support for anti-corruption platforms.

On the methodological side, there are limits to the conclusions the empirical results allow to draw. The online convenience samples, as with most survey experiments, might not be representative of the voter population in both countries. More importantly, even though the conjoint design is designed to approximate real-world voter decision making under trade-offs between two candidates differing across multiple dimensions, there are limits as to what we can infer from the conjoint experimental results. These limits are a function of design. In particular, we cannot infer that any candidate campaigning on an anti-corruption platform will *lose the election*, as Abramson, Koçak and Magazinnik (2019) have shown. In a similar vein, the external validity of the design is undermined by the uniform distribution of candidate characteristics, whereas in reality political candidates might have attributes that are

distributed non-uniformly (de la Cuesta et al., 2020), a problem which conjoint experiments on corruption are not exempt from (Incerti, 2020).

The design also does not allow to draw any final conclusions on the similarities and differences of the benefits of anti-corruption policies in low-corruption and high-corruption environments. More systematic research on this differential is urgently needed beyond the limited inferential strategy of simple case selection as employed in this paper. Finally, the limited sample size and non-experimental nature for AMCEs conditional on household income and partisanship prescribe an interpretation of these results as suggestive. More research is needed to address those limitations both substantive and methodological. Especially the conditioning of voter support for anti-corruption platforms by status as beneficiary of corruption is an exciting avenue for future research on corruption voting and electoral accountability.

Chapter 4

Tax Earmarking and Political Participation: Theory and Evidence from Ghana

4.1 Introduction

Building effective ‘fiscal states’ is often seen as the cornerstone of economic and political development (Schumpeter, 1918; Besley and Persson, 2009; Acemoglu and Robinson, 2019, 2022; Stasavage, 2020). Historically, the development of those fiscal states is thought to have followed a pattern by which particular taxes were first raised for specific purposes, for example earmarked for financing war efforts by early modern European rulers (Tilly, 1992; Kiser and Levi, 2015; Kiser and Karceski, 2017). According to these accounts, only with the development of modern welfare states were taxes slowly de-linked from particular benefits, and the notion of general consolidated revenue funds—made up of a wide range of different taxes to finance a wide range of public goods—emerged (Webber and Wildavsky, 1986; Levi and Kiser, 2015). Thus, how tightly taxes are linked to particular benefits has important implications for the relationship between citizens and the state more broadly.

In many countries today, the practice of earmarking (that is, raising and allocating) particular taxes for specific benefits remains widespread (McCleary, 1991; Wagner, 1991; Wilkinson, 1994; Hsiung, 2001). It is often promoted by international financial organizations to enhance fiscal transparency and sometimes even as preconditions for loan programs in many developing countries in the context of high levels of corruption and low levels of government accountability (Teja, 1988; IMF, 2007). Earmarked revenues have been estimated to typically account for between one-third and two-thirds of governmental expenditures in the United States (Wagner,

1991); in a 2018 OECD survey, all surveyed Latin American and Caribbean countries had earmarked funds, in some cases representing between 61% and 80% of all expenditures (OECD, 2020); and many governments in sub-Saharan Africa, most notably in Ghana, Uganda, and Côte d’Ivoire, have repeatedly resorted to tax earmarking, especially to finance the provision of health and education services (McCleary, 1991). In Kenya, for example, initial decentralization reforms earmarked five percent of the total national income tax revenue for the Local Authority Transfer Fund to increase funds available to local governments (Sheely, 2015). Other prominent examples include spending a fraction of revenues from the personal income tax on education or health, or earmarking motoring taxes for spending on local infrastructure.

However, recent literature has argued that earmarking can be an inefficient fiscal tool. Earmarked funds may cause budget rigidities due to the inability to fund programs that are in line with new or changing policy priorities (Wilkinson, 1994; Flores-Macías, 2018). In addition, some earmarked taxes resemble user charges and thus are constrained by the problem of truthfully revealing agents’ consumption of the good that is financed by earmarks or user charges (Gruber, 2019). It is unclear why earmarking remains a prevalent tool in public finance, being widely implemented in many developing tax states today while introducing inefficient budgetary rigidities. One way to rationalize this is to explore whether there is a decidedly political logic that has to do with a government’s fear of societal mobilization in response to general fund taxation. This chapter tries to think through the puzzling prevalence and persistence of tax earmarking and its political causes and consequences in many emerging “fiscal states” today. Why and under what conditions governments resort to tax earmarking remains understudied.

I focus in this chapter on the demand side, that is, the political consequences of tax earmarking, and in particular, society’s political responses to the state’s attempt at increasing or reducing spending discretion. We first need to understand whether governments actually have reason to fear differential political mobilization under the two tax-benefit linkages. Thus, in this chapter, I focus on citizens’ incentives to influence the government to achieve a variety of goals—that is, to exert voice—and how they might differ in response to different strengths of the tax-benefit linkage.

Although existing scholarship suggests that taxation enhances the will or ability of citizens to become politically engaged and hold government accountable (Ross, 2001; Paler, 2013; Martin, 2016; Weigel, 2020), we have little systematic theoretical and empirical work that explores whether this participation dividend of taxation might be different under different types of linking taxes with benefits. Indeed, under some

circumstances, tax earmarking might lead to greater citizen demands for government accountability due to improved transparency of the tax price of public goods (Dhillon and Perroni, 2001; Meadowcroft, 2011; Prichard, 2009, 2010, 2015; van den Boogaard et al., 2022). The argument is that tax earmarking “can give taxpayers a say over how tax revenue is spent, improve monitoring expenditures, build trust around taxation and encourage public engagement” (Prichard, 2010, p. 34). Extending this logic, Mosley and Abdulai (2020) argue that earmarking embodies the contractual, explicit side of fiscal contracts and thus strengthens the accountability relationship between citizens and the state. Finally, this expectation emerges as a corollary of recent scholarship that emphasizes an “ownership” mechanism through which taxation produces accountability pressures (de la Cuesta et al., 2021).

However, we have little systematic theoretical and empirical evidence on the accountability consequences of tax earmarking. In particular, it is unclear why governments would voluntarily strengthen the fiscal contract by earmarking tax revenue if it means greater citizen mobilization. In this chapter, I contrast the different competing predictions about the political consequences of tax earmarking that follow from (i) the literature on taxation and accountability and (ii) a simple decision-theoretic framework, lay out how we would know which predictions are correct, and subject them to a preliminary empirical test.

To reduce complexity owing to multiple parameters involved (citizens, governments, state capacity), I provide a simple decision-theoretic model of the incentives for citizens to engage in costly political action in response to the government’s decision to earmark tax revenue or not, yielding results that show the counterproductive role of providing information via earmarking in affecting citizen engagement. This model states that the vagueness of government commitments under general fund taxation increases government’s discretion over spending, which in the context of many low-capacity states with high levels of informality would mean greater incentives for rent extraction and coercive taxation (Paler et al., 2017; Henn et al., 2022). Thus, the signal that general fund taxation sends in many low-capacity states, where beliefs about political efficacy might be more volatile and vulnerable to shocks, is one of potential rent-seeking and corrupt behavior on the part of government. In response, citizens have incentives to (i) seek information to determine if spending actually misaligned with their preferences and (ii) exert voice to reduce future government discretion. These incentives are much weaker under earmarking as it provides precise information on the tax-benefit linkage.¹

¹Note that this type of societal mobilization is much broader and more general in nature than mere

Thus, the main contribution of this chapter is theoretical rather than empirical. I provide a theoretical clarification of the relationship between the tax-benefit linkage and political participation. While most of the existing literature thinks that earmarking taxes has positive consequences for the demand for accountability, by formalizing intuitions and clarifying mechanisms as well as assumptions we can reach conclusions opposite to the current way of thinking about the political consequences of earmarking. The framework suggests a set of important parameters that matter for the relationship between tax earmarking and political participation, and it explores which values of these parameters yield which conclusions. Consequently, what I offer in this chapter is a set of intuitions that yield multiple observable and testable implications for future empirical research to embrace with more sophisticated empirical tools than the ones employed in this chapter.

Thus, this chapter provides only a first attempt at testing these competing claims empirically. I conducted an online survey experiment with 874 respondents in Ghana, a country in which tax earmarking has featured prominently in policy debates and citizen engagement in the past decades (Prichard, 2009, 2010, 2015). I randomly vary different proposals of how to use increased tax revenue from a government fiscal capacity building program, that is, whether tax revenues are used to finance a particular public service or whether they contribute to the general government fund. Citizen participation is measured via a series of endline questions that tap into intentions for political engagement. Evidence on the proposed mechanism comes from secondary outcome measures of discontent with the lack of government-provided fiscal information and fear of increased government discretion over spending. I present preliminary evidence that tax earmarking makes no difference to citizens' propensity to participate in political life. Two explanations for this null result emerge from the data: citizens may have lacked belief in (i) general government responsiveness to citizen efforts at influencing political life, and (ii) the collective action capacity of society as a whole to overcome coordination problems surrounding political participation. These results suggest that we may need to revise our (positive or negative) view of tax earmarking and its implications for political participation. However, another explanation may well be that the experimental environment was simply too artificial to capture any underlying real-world cost and benefit analyses that citizens may make in response to earmarking. Future work is needed to test more thoroughly the set of intuitions

support for different tax-benefit linkages (Kallbekken and Aasen, 2010; Brockmann, Genschel and Seelkopf, 2016; Flores-Macías, 2018). Thus, the type of civic engagement examined in this chapter has much broader underlying mechanisms than mere tax resistance in the form of reduced tax morale.

and observable implications that follow from the theoretical framework proposed in this chapter.

4.2 Background and Motivating Example

How strongly taxes are linked to particular benefits is defined by the concept of the tax-benefit linkage (Gruber, 2019). On one end of the spectrum, *earmarked taxation* (or hypothecation) denotes “the practice of designating or dedicating specific revenues to the financing of specific public services” (Buchanan, 1963, p. 457) or, alternatively, “the practice of assigning revenue [...] from specific taxes or group of taxes to specific government activities or areas of activity” (McCleary, 1991, p. 82). Thus, revenues from a specific tax are earmarked for a specific spending item. This makes clear exactly which benefits are associated with a given tax (increase). On the other end of the spectrum, under *general taxation*, tax revenues extracted from citizens makes up a consolidated (or general) revenue fund; thus, general taxation is paid into the consolidated fund, and general spending is paid out of the consolidated fund (Levi and Kiser, 2015; Kiser and Karceski, 2017). Here, the relationship between taxes and spending is loose and resembles a general fiscal exchange (Martin, Mehrotra and Prasad, 2009). In this case, a given tax (increase) typically finances a wide variety of budgetary spending items.

Thus, one of the crucial differences between earmarking and general taxation that may matter for political behavior is that under earmarking, the immediate benefit of any given tax (e.g., a public good) is known to the taxpayer, whereas under general taxation, this information is much more vague since citizens may know the tax costs of each tax they already pay yet, under general taxation, they do not know which of these taxes (i.e., in terms of the tax base) is used to finance a particular public good in question; hence, the benefit associated with a given tax (or an increase in the rate of this tax) is unknown to taxpayers.² Under general taxation, citizens do not know the revenue purpose of a particular tax (or an increase of this tax) even though they might have the same overall tax burden as they would have under earmarking. Thus, they do not know its immediate tax benefit. Crucially, this vagueness of government spending commitments also increases government’s discretion over spending, which

²In some circumstances, earmarking might effectively be achieved through partisan politics and citizens’ voting for particular programmatic appeals that “ earmark” revenue for policies that are driven by the ideological orientations of political parties. However, in contexts of high levels of informality, demand for programmatic policy is drastically reduced (Gottlieb, 2022).

in the context of many low-capacity states would mean greater incentives for rent extraction or even corruption on behalf of the government.³

4.3 Theoretical Framework

4.3.1 Preferences

The main contribution of this chapter is to provide conceptual clarity on how to think about the effect of earmarking on political participation. What would citizen engagement in response to these financing instruments look like? Without this kind of model, we would not know which (alternative) mechanisms to test for or which sources of heterogeneity to account for. In what follows, I model one way to think about the politics of tax earmarking to derive testable and competing implications about the role of the tax-benefit linkage in structuring political participation. This framework explores important parameters in the relationship between earmarking and participation, and quantifies under which conditions these parameters yield a given conclusion. I thus provide a set of intuitions which can then be fruitfully explored by future research with more sophisticated empirical tools.

We model a world in which society consists of a continuum of citizens of size $a > 0$, indexed by $C \in [0, a]$.⁴ The government announces a tax increase where this tax instrument can either be tax earmarking for a particular public good or contribute to the general consolidated revenue fund so that $T \in \{e, g\}$. The crucial insight here is that we precisely do not assume the government to make this choice exogenously, but endogenously to the trade-off between signalling its type and securing greater

³Earmarked taxation can take on a wide variety of different forms and different typologies (McCleary, 1991; Wilkinson, 1994). For analytical clarity, we consider here the case of what is known as *strong and wide* earmarking, denoting the designation of tax revenue to a particular spending purpose where (i) the amount of revenue from the tax determines the amount of spending on the service (*strong*) and (ii) revenue is designated for a whole spending function, for example healthcare (*wide*). In either case, what varies between earmarking and general taxation then is the amount of information provided about those tax costs since in the earmarking case, a specific tax is tied to a particular end use.

⁴We assume that there is some non-zero baseline level of political participation, that is, citizens are generally able to overcome collective action dilemmas some of the time, for example via the provision of private benefits (Olson, 1971), psychological rewards (Bueno de Mesquita and Shadmehr, 2022) or reduced social costs (Paler, Marshall and Atallah, 2018). This may be achieved through taxation itself as a growing literature stipulates that the act of taxing citizens makes them politically more active (Ross, 2001; Paler, 2013; Weigel, 2020). Thus, it may be possible that taxation—no matter the tax-benefit linkage—induces higher levels of participation. The question with which this chapter is concerned, however, is whether the tax-benefit linkage differentially shapes this baseline rate of participation.

policy discretion over spending. This gives citizens varying incentives to lobby the government to achieve various different goals.

C can in response lobby the government to obtain a variety of different goals, l , so that the action set is $A_C = \{l, \neg l\}$ with a cost associated with lobbying, c_l . This can be thought of as exerting voice (Hirschman, 1972) in the classical sense of political participation and collective action; it may also include activism in interest groups. The benefits from lobbying B_l can represent one of three motivations: greater alignment of the public good provided with individual policy preferences (in the form of the Euclidean distance from the citizen's ideal point, $B_l \equiv -|x - z_C|$); greater information and transparency about spending outcomes for tax revenue; and a reduction of future government discretion over spending tax revenue. Those benefits are obtained with probability $p_e \in [0, 1]$ and $p_g \in [0, 1]$ which represent the probabilities that citizen lobbying efforts are successful.⁵ Thus, expected utility from lobbying in response to earmarking is

$$EU(l) = p_e B_l - c_l, \quad (4.1)$$

and expected utility from lobbying in response to general fund taxation is

$$EU(l) = 2p_g B_l - c_l. \quad (4.2)$$

The different specifications of these utility functions reflect the different types of incentives that citizens face in response to an earmarked or non-earmarked tax increase, an intuition we will further develop now.

4.3.2 Incentives

Our main task is to determine the different incentive structures for citizens to lobby the government in response to the two financing instruments.

$T = e$. Under earmarking, government releases perfect information on the specific revenue purpose of the tax increase so that citizens learn exactly whether they are net beneficiaries or net losers from the tax-and-spend policy, that is, the size of B_l . The benefits from lobbying under earmarking stem purely from the fact that lobbying to

⁵This allows for the tax-financing instrument to not only affect benefits from lobbying B_l , but also for negative or positive shocks to citizens' beliefs about the probability of successful lobbying when subjected either to an earmarked or non-earmarked tax increase. Alternatively, this probability can be thought of as citizens' perceptions of (i) political efficacy and responsiveness of the system, as well as (ii) the credibility of any commitment or announcement the government makes.

change the earmarked policy becomes more beneficial the further away that policy is from an individual's ideal point, i.e., the greater B_l is.⁶ Consequently, we have as the condition for lobbying in response to earmarked taxation that

$$p_e B_l - c_l > 0. \quad (4.3)$$

Clearly, lobbying is only beneficial if and only if $p_e B_l > c_l$. Thus, citizens for whom the earmarked policy x already represents their ideal point z_C , i.e., when $B_l = 0$, will have no incentive to lobby and choose not to lobby for policy change. By contrast, and assuming earmarking induces some non-zero belief in the success of lobbying $p_e > 0$ (that is, the efficacy of political participation), only citizens for which $B_l > 0$ and the costs of lobbying are sufficiently small have incentives to lobby the government in response to earmarked taxation. Further notice here that this inequality is true when $p_e > \frac{c_l}{B_l}$. Thus, p_e has to be quite high in order for expected utility from lobbying to be positive, that is, larger than the cost-benefit ratio of participation. For example, when the costs of lobbying are half of the benefits from lobbying, then the inequality is only true when $p_e > \frac{1}{2}$, that is, citizens belief that half of the time their lobby attempts at changing the earmarked policy will be successful.

Citizens lobby the government in response to earmarked taxation whenever

- a. the benefits from lobbying for policy change (B_l) are sufficiently high,
- b. the costs of lobbying (c_l) are sufficiently low.
- c. beliefs about government responsiveness (p_e) are sufficiently high.

$T = g$. Under general fund taxation, the tax financing instrument is vague, i.e., the tax increase is not directly tied to financing any particular public good but instead more loosely related to the general consolidated revenue fund. Thus, citizens do not learn which exact public good(s) are financed by any given tax increase. Crucially, this vagueness of government commitments also increases government's discretion over spending, which in the context of many low-capacity states would mean greater incentives for rent extraction or even corruption on behalf of the government. Thus, the signal that general fund taxation sends in many low-capacity states is one of potential rent-seeking and corrupt behavior on the part of government. In response, citizens have incentives to (i) seek information to determine if spending is actually

⁶There are no incentives to lobby the government to obtain information about the revenue target of the tax increase since that information is already provided; nor to reduce future government discretion since that has already been reduced by virtue of earmarking.

misaligned with their preferences and (ii) exert voice to reduce future government discretion.⁷

Amount and nature of the benefits from lobbying the government in response to general taxation are thus different in two ways:

1. Citizens receive B_l twice if they choose to lobby the government in response to general fund taxation, that is, every citizen (and not just policy net losers as in earmarking) has two incentives to mobilize. First, they receive benefits from seeking information to find out whether *current* revenue is spent in line with their preferences, that is, exactly where the tax money extracted today will go. Second, by choosing not to earmark the tax, the government signals its preference for retaining discretion over spending in the *future*, an outcome that is costly for citizens in the form of forgone utility from alignment of public policy with individual preferences in future distributions of spending. Thus, citizens receive benefits from exerting voice to politically constrain the government in its spending discretion in the future, e.g., via pushing for earmarking. Thus, preferences for earmarking stem from a desire to obtain spending information from the government necessary for accountability pressures and to reduce spending discretion of future governments.
2. However, these benefits are partially offset by (i) the costs of lobbying and (ii) the probability that these lobbying efforts will be successful, p_g . Here, we allow for the possibility that the choice of general fund taxation might not just positively affect B_l but also negatively affect p_g should citizens update about more *general* political efficacy and the political systems responsiveness to citizen lobbying efforts. The crucial insight, which we will demonstrate below, is that even with this constraint, general fund taxation might produce more incentives

⁷Of course, government discretion over spending can be beneficial for co-partisans, co-ethnics or clients within clientelist networks. In this case, discretion is what *enables* the distribution of club goods. However, in this model, even those who benefit somewhat from government discretion have a countervailing incentive to increase overall responsiveness of the government in future periods since they cannot be sure that their party (or ethnicity) retains political power, especially in contexts of high political uncertainty. This is especially true when examining the politics of earmarking for providing *public* goods that cannot be handed out selectively to particular winning coalitions. Thus, together with the fact that different types of taxes impose different burdens on different segments of society, there will be a distribution of net policy winners and losers under general taxation as well, yet the main intuitions of the argument remain unaffected by this possibility as we assume the type of tax to be uncorrelated with its tax-benefit linkage.

to lobby the government assuming that citizens do negatively update, but not drastically so.⁸

In this case, the condition for lobbying in response to a non-earmarked tax increase is that

$$2p_g B_l - c_l > 0. \quad (4.4)$$

Clearly, this is again true when the expected benefits from lobbying are larger than the cost of lobbying such that $2p_g B_l > c_l$. How small can the belief in general government responsiveness be and still yield positive expected net benefits from lobbying? In other words, how large does the scepticism about the effectiveness of lobbying in citizens beliefs have to be in order for citizens to decide not to lobby in response to general taxation? It turns out, quite large. Precisely, in order for Eq. 4.4 to still hold, $p_g > \frac{1}{2} \frac{c_l}{B_l}$. Thus, whenever p_g is sufficiently small, citizens lobby in response to general fund taxation. For example, when the costs of lobbying are half of the benefits from lobbying, then it only need to be true that $p_g > \frac{1}{4}$, a dramatically lower belief that the government is responsive and lobbying will be successful compared to the earmarking case. This is also the condition for comparing levels of participation under general taxation to policy beneficiaries of the earmarked good ($B_l = 0$). When general taxation leads to a reduction in the general beliefs about political efficacy down to $p_g \leq \frac{1}{4}$, it will not produce greater participation than earmarking even when compared to those who are policy beneficiaries from the earmarked good. As such, citizens' beliefs about general responsiveness p are different from the fear of future government discretion over spending policy which are the basis for the benefits of lobbying under general taxation B_l .

Citizens lobby the government in response to general fund taxation whenever

- a. the benefits from lobbying for policy change (B_l) are sufficiently high,
- b. the costs of lobbying (c_l) are sufficiently low,
- c. beliefs about government responsiveness (p_g) are sufficiently high.

⁸The intuition for why general fund taxation makes citizens fear increased government discretion but not necessarily dramatically reduced political efficacy is that the latter concept has been shown to be the rather deeply ingrained outcome of more long-run processes of socialization, formed over years of experience with government (Lieberman and Zhou, 2021). Thus, a one-shot implementation of general fund taxation might not be very informative about general government responsiveness.

Besides quantifying the conditions under which citizens lobby in response to *either* tax-financing instrument, the more interesting question is when one instrument produces more participation than the other. Thus, we now quantify precisely under what conditions general fund taxation produces more political participation than earmarking when $B_l > 0$ under earmarking, that is, compared to citizens for whom the earmarked good is not the ideal point. Put differently, when will

$$2p_g B_l - c_l > p_e B_l - c_l. \quad (4.5)$$

Clearly, this is true when the expected utility from lobbying under general taxation are greater than the expected utility from lobbying under earmarking, $2p_g B_l > p_e B_l$ (or $p_g B_l > \frac{p_e B_l}{2}$). How large can the negative shock to citizens' beliefs about systemic political efficacy be under general taxation in order to still obtain greater levels of participation (even compared to those whose ideal point is not congruent with the earmarked good)? Assuming $B_l > 0$ for citizens under earmarking, inequality 4.5 is true whenever $p_g > \frac{p_e}{2}$, that is, the ratio of the two probabilities of systemic efficacy needs to be less than 1 : 2. For example, if earmarking has increased the prior $p_e = 0.5$ to $p_e = 0.75$, the negative shock to these beliefs under general taxation cannot lead to a decrease down to less than $p_g = \frac{0.75}{2} = 0.35$. Whenever general taxation leads to a general reduction in the overall belief that systemic political efficacy that is much lower than this (even though there might be possible benefits from trying to reduce government spending discretion), earmarking will produce greater participation among those that are net policy losers.⁹

In short, under certain conditions, general taxation will still yield greater benefits from lobbying even when compared to those in earmarking who stand to gain the most in terms of policy change, that is, net policy losers. This is because the incentives to lobby under general taxation are qualitatively different *and* twice as large as those under earmarking. At the extreme, when $B_l = 0$ in earmarking, citizens will choose

⁹Note that this condition becomes less strict if we allow for another parameter in the utility function from lobbying in response to earmarking that captures the probability that in fact $B_l > 0$. This would represent the probability of the ideal point being incongruent with the earmarked policy, against which the benefits from lobbying would additionally need to be weighed against. We omit this additional complication here for simplicity but note that introducing this additional constraint would yield even stronger predictions about the participation dividend of general fund taxation as is evident from analyzing the case of $B_l = 0$ under earmarking as conducted above. This implies that we keep the distribution of preferences over policy constant across strengths of the tax-benefit linkage so that preferring a particular kind of policy is unrelated to the tax-benefit linkage.

not to lobby and thus the comparison of expected utilities is equivalent to the one in Eq. 4.4.¹⁰

Citizens lobby the government in response to general fund taxation more than in response to earmarking whenever the reduction in beliefs about systemic political efficacy in response to general taxation is small enough, that is, as long as $p_g > \frac{p_e}{2}$. Note that if this is not the case, that is, when earmarking simultaneously increases beliefs about the government's responsiveness p_e sufficiently, policy change becomes a central motivation to engage in politics and may increase in relative importance vis-à-vis information seeking and discretion reduction. This is because the likelihood of policy change, conditional on lobbying, increases and thus there is a chance that changing earmarking's target would be effective. Thus, the pre-existing distribution of policy preferences would matter a lot in this case yet it would be much less relevant under general fund taxation since the policy is unknown ex ante. This would change the notion of how strong a commitment earmarking really is, and it would imply that it may be more costly for the government to switch from earmarked item 1 to earmarked item 2 than it is to switch from general fund taxation to earmarked taxation. Thus, varying beliefs about government responsiveness induced by the tax-benefit linkage itself may produce competing conclusions.

This theoretical framework formalizes three main intuitions.

1. Earmarking and general fund taxation produce different incentives for citizens to engage in costly political action with different motivations for participation. These incentives arise purely due to the differential amount of information provided by the government about the spending purposes of taxation. The signal that general fund taxation sends in many low-capacity states is one of potential rent-seeking and corrupt behavior on the part of government. In response, citizens have incentives to (i) seek information about *current* government spending to determine if spending is actually misaligned with their preferences and (ii) exert voice to reduce *future* spending discretion the government would enjoy.

¹⁰A source of heterogeneity not captured in the formal argument above is citizens' prior experience with public goods provision (which may or may not be orthogonal to beliefs about political efficacy). When public goods provision by the formal state is extensive and broad, citizens are more willing to pay taxes (Bodea and LeBas, 2016; Castañeda, Doyle and Schwartz, 2020), making tax evasion as an exit option less and exerting voice more likely. Prior experience with public goods provision may also alter which signals different tax-benefit linkages would send in these contexts. Earmarking tax revenue for successful public goods provision reduces incentives to mobilize politically, whereas general fund taxation in a context of broad-based public goods provision might not be interpreted as a discretion-enhancing strategy by the government, reducing incentives to mobilize. Overall, this line of reasoning does not predict a *difference* in participation rates across different tax-benefit linkages, just that overall levels of participation might be lower.

These incentives will need to be balanced against the probability of government responsiveness to these lobbying efforts which may be lower under general taxation; yet, as long as these beliefs are not completely zero, citizen lobbying in response to general taxation should be greater than to earmarked taxation, and it is driven by different motives. In other words, if the negative shock to general political efficacy from general taxation is not too large, general fund taxation should produce higher levels of participation by a broader segment of society and that goes beyond simply lobbying for policy change.

2. Although citizen voice should be lower on average in response to earmarked taxation than to general taxation, we still expect lobbying to be beneficial in response to earmarked taxation for the subset of citizens for whom the current earmarked good is not the ideal point as they have an incentive to lobby for policy change. By contrast, *all* citizens have two incentives to engage in political action in response to general taxation: they seek information in the absence of government-provided transparency, and they lobby for greater accountability as they infer from the government's decision not to earmark that it tries to increase its future spending discretion and thus rent-seeking. However, the critical difference here is that only for some citizens is $B_i > 0$ under earmarking, whereas it is $B_i > 0$ for all citizens under general taxation since in that case, utility is not derived from ideal policy but from more precise policy information and greater government accountability. Thus, even for low values of beliefs in government responsiveness, general taxation should produce more citizen participation.
3. These implications are particularly visible for citizens whose ideal point is already implemented in the current earmarked policy, i.e., for those with $B_i = 0$ whose not to lobby instead and obtain a payoff of 0. In this case, even when beliefs about government responsiveness are very low, general taxation may still produce more citizen mobilization.

We can derive the following main empirical implication from the propositions above.

- H1 Citizens are more likely to engage in political action in response to general taxation than to earmarked taxation.

In addition, if the mechanisms outlined in the theoretical framework above are at play, we should observe that compared to earmarked taxation, citizens are more likely to engage in political action to seek revenue information (H3a) and reduce government discretion (H3b) in response to general fund taxation. By contrast, compared to general fund taxation, citizens are more likely to engage in political action to change policy (H3c) in response to earmarked taxation, especially when the benefits from the earmarked good are higher (H3d). We should also observe empirically that compared to earmarked taxation, the increase in participation in response to general fund taxation is higher the higher the beliefs about government responsiveness (H4a) and the lower the benefits from the earmarked good (H4b).

4.3.3 Main Assumptions

With this general setup, we can state the main assumptions of this model.

1. Citizens know that only high-type, responsive governments choose to earmark increased tax revenue, and only low-type governments choose to not earmark revenue. Thus, the tax-benefit linkage is a signal about government type p which then affects whether citizens get additional benefits from trying to reduce the government's desired discretion over spending, B_l . By implication, we assume that citizens do not interpret earmarking as a cheap-talk transparency initiative designed by a low-type government to increase its popularity. This is intuitive since low-type governments potentially have larger benefits from increasing their discretion over spending so enable spending on private benefits for a narrow selectorate and winning coalition.
2. The tax-benefit linkage affects short-term fluctuating beliefs about government's discretion over spending (which constitutes a part of B_l) more strongly than long-term beliefs about the general responsiveness of the political system to citizens' lobbying efforts (p). This means that the decision to not earmark tax revenue is a stronger signal about potential benefits from lobbying the government to reduce its discretion over spending than it is about general political efficacy. This is believable in light of a body of literature showing that political efficacy is a rather deeply ingrained outcome of more long-run processes of socialization, formed over years of experience with government (Lieberman and Zhou, 2021). Thus, values of p might be low to begin with in many contexts but that is unlikely to be due to short-term changes in tax-financing instruments. Rather, a legacy of strong democratic norms, high levels of public contestation

and protests as well as freedom of information laws combined with competitive elections might shape feelings of political efficacy and the effectiveness of lobbying more strongly (Ashworth, 2012; Kosack and Fung, 2014; Grossman, Michelitch and Prato, 2022; Grossman and Slough, 2022). This may also travel to the African context where we do observe citizen accountability pressures via competitive elections (Harding, 2020); in fact, norms of ruler contestation and claim-making in this setting go back to pre-colonial institutions of autonomous local communities which constrained precolonial rulers (Bolt et al., 2022).

3. Furthermore, there are substantial costs to the government of citizen mobilization that include the revelation of information the government does not wish to reveal, change in policy the government does not wish to change, or reduced electoral prospects. We assume here that citizens are aware of these substantial political costs and that this in part drives their baseline beliefs about p . However, there is enough information uncertainty about likely societal contestation that we can, in this chapter, exogenize the choice of the tax-financing instrument and examine political responses to it, regardless of citizens beliefs about the likely political costs of earmarked and general fund taxation.
4. All three motivations for citizen lobbying are equally valued by citizens so that they can be compared as B_l or multiples of B_l such as $2B_l$ as in the model. Getting current policy to align with individual preferences is as valuable for citizens as obtaining information about government spending of tax revenue and reducing the government's future spending discretion. This seems plausible given the central role of information in holding governments accountable for policy outcomes (Dunning et al., 2019). Especially in low-information environments, where uncertainty is prevalent, the benefits from obtaining information can be substantive and as large as the benefits from preference alignment or discretion reduction. Similarly, discretion reduction as a medium- to long-term goal is valuable if citizens' time horizons are sufficiently long. However, reducing the scope for discretion can be immediately relevant for citizens, for example the next time the government announces a new policy. Thus, even with short time horizons, accountability can be achieved (Ashworth, 2012).
5. Under $T = g$, citizens can obtain information about government spending *and* reduce future government discretion with a *single* form of lobbying (incurring only c_l instead of $2c_l$). This is believable, e.g., in the form of writing a letter or

signing a CSO petition that seeks to address both issues simultaneously. This also implies that information-seeking and lobbying for reduced discretion often happen simultaneously and not sequentially (so that the decision by the government to respond to information-seeking or not does not factor into citizens' choice).

6. We assume that seeking information is valuable to citizens regardless of whether they think the government will actually provide such information or has incentives to strategically release information in order to fend off potentially politically costly protests. This is consistent with there being some baseline political efficacy as well as a desire for information about what happens to taxpayers' money.
7. We assume that citizens possess sufficient information to determine alignment of their preferences with government policy. In addition, beliefs about government policy might be conditioned by an individual's previous experience with service delivery in low-capacity settings (Castañeda, Doyle and Schwartz, 2020). While this might be true in many contexts, *prima facie* it is relatively easy for citizens to determine whether the *announced* government policy agrees with one's preferences, and that is all the model needs to assume. In addition, the empirical case of this chapter, Ghana, is a setting in which government effectiveness is significantly higher than in all its neighboring countries according to the Worldwide Governance Indicators.¹¹

4.3.4 Alternative Mechanisms

The government's decision to earmark increased tax revenue or not can send signals different from the ones outlined above, possibly reversing the expected relationships between the strength of the tax-benefit linkage and citizen political participation as described in the previous section. Thus, there are mechanisms that can give rise to an alternative hypothesis, namely that citizens are *less* likely to engage in political action in response to general taxation than to earmarked taxation (H2). There are also two alternative mechanisms that lead to predictions similar to the ones obtained in the fear-of-discretion framework, that is, there might be greater civic mobilization in response to general fund taxation to reign in the coercive capacity of the state but for reasons different from the ones outlined in the theoretical framework.

¹¹<https://info.worldbank.org/governance/wgi/Home/Reports>.

1. It could be that, in response to general fund taxation, the magnitude of citizens' negative update about general political efficacy might be larger than the perceived rewards from seeking spending information and reducing the government's discretion and rent-seeking behavior. Thus, citizens' negative update of responsiveness beliefs in response to general fund taxation is greater than their positive update of beliefs in response to earmarked taxation (H5a), leading to higher or similar levels of mobilization under earmarking.
2. Earmarking might send the signal that the government has the power and discretion to decisively and confidently set the tax-benefit linkage and to unilaterally determine the earmarked good without further consultation by civil society and political efficacy (Lieberman and Zhou, 2021). Citizens would then be more likely to feel left out of the policy deliberation process, strengthening their incentive to constrain government spending discretion (H5b).
3. Earmarking could also signal that the government is trying to shield certain policy items as private goods for a narrow selectorate from future policy reversal under different governments (Bueno de Mesquita et al., 2003; Bueno de Mesquita and Smith, 2011), a signal that would increase citizens' incentive to mobilize politically in order to reduce this form of government discretion over spending. Thus, citizens would be less likely to predict policy reversal of private goods in response to earmarked taxation than to general fund taxation (H5c).
4. Earmarking could increase citizens' skepticism of the government and its performance if they associate earmarking with loan conditions by international organizations such as the International Monetary Fund. This would reduce the overall evaluation of government's fiscal performance as earmarking could be seen as conforming to IMF loan conditions which are only set out when the IMF does not trust the recipient country enough to spend public revenue diligently. Thus, citizens would be less likely to evaluate government as fiscally competent in response to earmarked taxation than to general fund taxation (H5d).

Finally, two mechanisms generate a hypothesis that predicts no difference between tax earmarking and general fund taxation (H6a). First, it could be that citizens are less likely to express willingness to pay taxes (exit) in response to general fund taxation than to earmarked taxation (H6b). According to this logic, citizens exit in response to general fund taxation instead of mobilizing, thus producing a null

result. Second, it could be that citizens are not more likely to expect other citizens to get involved in politics in response to general fund taxation than to earmarked taxation (H6c). According to this logic, if general fund taxation would really lead to more participation, then heightened beliefs in collective action should accompany that effect.

4.4 Experimental Design and Implementation

4.4.1 Setting

The rest of this chapter proceeds with a first and preliminary attempt at testing these competing hypotheses. An online survey experiment was conducted with 874 citizens in Ghana recruited via the survey firm Bilendi; this sampling strategy primarily recruited an online convenience sample of young and more educated segments of Ghanaian society, ones who are likely to be mobilized more easily (see Table C.1). The vignettes were intended to manipulate respondents' perceptions of the extent of tax earmarking which is essentially a continuum of "varieties of earmarking" when applying the standard public finance taxonomy of hypothecation (Wilkinson, 1994). In order to avoid deception of participants, I constructed the vignettes using real-world policy objectives of the current Ghanaian government as well as present *possible* options for how to use tax revenue in a hypothetical case, that is, options the government *can* consider.¹²

Ghana is an ideal case for examining the politics of tax earmarking for several reasons. A prominent example of the political struggle about general taxation and earmarking comes from Ghana (Prichard, 2009, 2010, 2015). In 2000, public opposition erupted in response to the government's announcement to increase the VAT rate from 10 to 12.5%, an increase in general tax revenue that would feed into the consolidated revenue fund. Appeasement was only achieved through earmarking revenues from the VAT increase to the newly established Ghana Education Trust Fund. The same public response in the form of mobilization ensued again in 2003 when a VAT rate increase to 15% was announced without further earmarking. In response to citizen contestation, the tax increase was renamed the National Health Insurance Levy and allocated to the creation of the National Health Insurance Fund. In March 2022, public opposition to the planned "e-levy", a highly contested tax on electronic

¹²The pre-analysis plan for this design was pre-registered at <https://osf.io/7a5m4>. IRB approval was obtained from the University of Oxford Departmental Research Ethics Committee (DREC), Ref. No.: SSH_DPIR_C1A.23_047.

transactions designated for the consolidated revenue fund, was successful in lobbying for a rate reduction from an initial 1.75% to 1.5% (Anyidoho et al., 2022). In this context, the politics of earmarked taxation remain highly salient and therefore constitute an ideal case for studying the accountability implications of the tax-benefit linkage.

4.4.2 Treatments

Respondents were randomly assigned to read two different vignettes intended to prime them on the extent to which government revenue is earmarked. Respondents in all treatment groups were presented with informational vignettes that start with an introductory tax-and-spend prime: “The government of Ghana has currently two important policy priorities: 1) improving national public health in the country which would benefit people like you, and 2) increasing the collection of tax revenues from a broad set of different taxes. The government is currently considering what to do with this additional tax money.” The general revenue fund treatment, that is, the control group, then read the following text in addition to the introductory prime: “One option the government can consider is that the money raised through the *extra taxes would go to the general revenue fund managed at the discretion of the government.*” The earmarking treatment group read the following text in addition to the introductory prime: “One option the government can consider is that the money raised through the *extra taxes would go only to a dedicated government revenue fund which is responsible specifically for these public health policies, the National Health Insurance Fund (NHIF).*” Treatment assignment was administered via Qualtrics. Treatment groups were balanced across a variety of pre-treatment covariates (Table C.1).

The introductory tax-and-spend text primes as policy issue public health for two reasons. First, it makes reference to the policy issue to make sure that the policy issue, in this case public health, is held constant across treatment groups. Second, public health was chosen because it prominently features in the history of the politics of tax earmarking in Ghana as described above, and because the tax-and-spend link is more credible than for other policy issues that are more prone to corruption, such as investment or infrastructure projects. The control group (general revenue fund) then primes a policy option that the government can consider: additional revenue from the tax campaign flows into the general revenue fund instead of being earmarked for a particular spending item. This is done to maximize the internal validity of the experiment, that is, to have a control group that resembles as closely as possible the concept of general fund taxation (the opposite of earmarking). Of course, this might come at

the expense of the external validity of the experiment since one could argue that in reality, no government can credibly commit to not earmark any new additional revenue due to possible political resistance. The treatment group (tax earmarking) then primes another policy option that the government can consider: additional revenue from the tax campaign is earmarked for the National Health Insurance Fund. This institution featured prominently in the tax episodes of the 2000s in Ghana mentioned above as the increase in VAT was earmarked for it directly.

Three points regarding the manipulation of the concepts of interest are worth noting. First, respondents spent on average 38 seconds on the treatment group vignette and 37 seconds on the control group vignette. Second, in a treatment-irrelevant Factual Manipulation Check (FMC), 67% of respondents correctly reported that the topic of the vignette was related to public health and not to digitalization or agricultural development. Third, in a series of treatment-relevant FMCs, participants' responses reported in Table C.2 suggest weakly that the earmarking and general revenue fund treatments successfully manipulated the underlying concepts of interest. Respondents in the treatment group were more likely to agree with the statement that taxes are closely linked to benefits, more likely to agree with the statement that the government transparently releases spending information, and less likely to agree with the statement that the government has a lot of policy discretion. However, these differences are not statistically significant.

4.4.3 Outcome Measures

I measure political participation in a variety of ways. First, respondents are asked: "For each of the following activities, please indicate whether you would engage in politics." Respondents can then indicate whether they would 1) talk to others in their neighborhood, 2) sign a petition, 3) contact their local MP, 4) attend a citizen assembly meeting, 5) join a civil society organization, 6) vote in upcoming elections, and 7) join a protest. The number of activities is then aggregated and counted on a scale [0, 7]. This is labelled *Political participation*. Second, this procedure is then repeated with slightly different question wordings that tap into the underlying mechanisms. These questions are modified to include the aim or motivation of political participation: "For each of the following activities, please indicate whether you would engage in politics to [get the government to release information on how it will spend taxpayers' money/reduce the amount of freedom the government will have in the future to decide what happens with your taxes/change how the increased tax revenue

from the government’s target is spent].” These are labelled *Information seeking*, *Discretion reduction*, and *Policy change*, respectively. Finally, respondents are asked an open-ended question: “Reflecting on your answers, why did you respond the way you did?”

It is important to bear in mind that this measurement strategy only elicits stated preferences as opposed to revealed preferences. Measuring behavior in the context of online survey experiments is challenging for a number of reasons. First, political behavior frequently takes place offline and in people’s immediate environment so that imitating this kind of environment online is difficult. Second, available alternatives to stated preference questions are sometimes difficult to justify from an ethical viewpoint. Encouraging people to follow hyperlinks to certain government websites, to have their names put on online petitions or to make donations to political organizations all come with their own practical, logistical and ethical challenges. Thus, this chapter taps more into the *latent* dimension of political participation, that is, an underlying hypothetical willingness to engage in politics more generally.

Quantitative text analysis of an open-ended question asked post-treatment confirms that the outcome measure meaningfully represented one way of expressing how and why citizens wanted to become active in political life. The question asked: “Reflecting on your answers, why did you respond the way you did?” The word “government” was by far the most commonly written word, followed by “citizens”, “country”, “taxes”, “politics”, “protest” and “right” (see Figure C.1). Topic modelling confirms this as is evident from Table C.4. An entire topic is devoted to words such as “government”, “taxes”, “know”, “used” and “accountable”. Another topic includes words such as “politics”, “vote”, “power”, “change” and “party”.

4.4.4 Analysis

To estimate the sample average treatment effect (SATE), I estimate the following OLS regression equation, calculating heteroskedasticity-robust standard errors:

$$Y_i = \beta_0 + \beta_1 K_i,$$

where K_i is a binary variable indicating one of the two experimental groups, and Y_i is the respective outcome variable for any given hypothesis test as in Table 4.1. All Likert-scale outcomes that measure agreement with a statement are transformed into an interval scale that ranges $[1, 5]$ where 5 indicates strong agreement.

In addition, I estimate the main SATE with covariate adjustment by controlling for predictive (of the outcome) pre-treatment covariates X_i ; these covariates are respondent age, gender, education, ethnicity, partisanship, as well as responses to survey questions that asked about their perception of government responsiveness and their preference for carrying out government policy through dedicated agencies. Heterogeneity in treatment effects is tested by interacting the treatment indicator with the moderating variable of interest as pre-registered in the pre-analysis plan; these are respondents perception of government responsiveness, their perceptions of policy benefits from earmarking, as well as respondent co-ethnicity with the main ethnic group, co-partisanship with the incumbent president’s party, and gender.

For the estimation of the main SATE, several additional p-values are calculated such as randomization inference p-values, as well as p-values adjusted for multiple comparisons: the Holm correction to control the False Discovery Rate (FDR) and the Benjamini-Hochberg correction to control the Family-Wise Error Rate (FWER). The open-ended question is analyzed by estimating LDA topic models and depicting the most common words written. A pre-registered power analysis confirmed that for the simple difference-in-means analysis of the effect of the treatment on the primary outcome variable Y_i , with 999 respondents I would be able to detect an effect of 0.05 points on the scale $[0, 7]$ given a standard deviation on the outcome variable of 0.282. Post-hoc power analysis with an empirically determined standard deviation of the outcome variable of 1.99 and the adjusted sample size reveals that the design was powered enough to detect an effect size of 0.38.

4.5 Results

4.5.1 Main Results

The estimates of the main sample average treatment effect are presented in Table 4.2. Column 1 shows the effect of tax earmarking on political participation estimated via OLS. There are no statistically significant differences across treatments in respondents’ propensity to engage in politics. The tax earmarking treatment increased the number of activities that respondents would engage in politically by only 0.007 points on a scale $[0, 7]$ and with a control group mean of 4.22. Several additional p-value calculations (randomization inference, multiple hypothesis testing corrections) confirm this null result. Additionally, this result is robust to covariate adjustment as reported in Table C.3. Thus, it appears that tax earmarking does not reduce citizens’ incentive to engage in political life as hypothesized above (H1). At the same

Table 4.1: Hypotheses Tests and Predictions of Treatment Effects

Hypothesis	Outcome	Prediction
Panel A: Primary hypothesis		
H1	Y_l	$K_g > K_e$
H2	Y_l	$K_g < K_e$
Panel B: Mechanism		
H3a	$Y_{l:I}$	$K_g > K_e$
H3b	$Y_{l:D}$	$K_g > K_e$
H3c	$Y_{l:P}$	$K_e > K_g$
H3d	$Y_{l:P}$	$K_{e,B_l>0} > K_{e,B_l=0}$
Panel C: Heterogeneity		
H4a	Y_l	$K_{g,\theta 1>0} - K_{e,\theta 1>0} > K_{g,\theta 1\gg 0} - K_{e,\theta 1\gg 0}$
H4b	Y_l	$K_g - K_{e,B_l=0} > K_g - K_{e,B_l>0}$
Panel D: Alternatives		
H5a	$ \Delta Y_\theta $	$K_e > K_g$
H5b	Y_L	$K_e > K_g$
H5c	Y_R	$K_g > K_e$
H5d	Y_M	$K_e > K_g$
Panel E: Null results		
H6a	Y_l	$K_g = K_e$
H6b	Y_C	$K_e > K_g$
H6c	Y_E	$K_g = K_e$
Panel F: Manipulation checks		
	$Y_{FMC-TR1}$	$K_e > K_g$
	$Y_{FMC-TR2}$	$K_e > K_g$
	$Y_{FMC-TR3}$	$K_g > K_e$

Notes: Y_l is the vector of individual lobbying decisions in response to the treatment, indicating general citizen participation. Subscripts $\{I, D, P\}$ indicate lobbying for information (I), to reduce discretion (D) and to change earmarked policy (P). In addition, $\Delta Y_\theta \equiv Y_{\theta 2} - Y_{\theta 1}$ which simply denotes the change score from endline to baseline of respondents' beliefs about government responsiveness Y_θ . Subscripts $\{L, R, M, C, E\}$ denote endline survey measures of citizens' perceptions of being left out of the policy process (L), of the likelihood of policy reversal (R), of the government's management of public finances (M), the reported tax morale measure (C), and expectations about others' involvement in politics (E). Outcomes $\{Y_{FMC-TR1}, Y_{FMC-TR2}, Y_{FMC-TR3}\}$ denote factual manipulation checks that are specific to the treatment groups.

time, it does not increase participation and accountability demands as hypothesized by a growing literature on the political economy of taxation (H2). This null result is therefore counterintuitive given the previous expectations of existing literature on the relationship between taxation and accountability.

Table 4.2: Effect of Tax Earmarking on Political Participation

	Participation	Information seeking	Discretion reduction	Policy change
Tax earmarking	0.007 (0.150)	-0.044 (0.153)	-0.114 (0.154)	-0.189 (0.149)
Control mean	4.220	4.882	5.013	5.213
RI p-value	0.970	0.773	0.446	0.237
Holm p-value	0.962	0.774	0.461	0.206
Benjamini-Hochberg p-value	0.962	0.774	0.461	0.206
Num. obs.	703	730	740	745

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Coefficients on *Tax earmarking* are OLS estimates with heteroskedasticity-robust standard errors. *Participation* is the number of political activities respondents indicated they would engage in. *Information seeking* is the number of political activities respondents indicated they would engage in to seek policy information. *Discretion reduction* is the number of political activities respondents indicated they would engage in to reduce the government's policy discretion. *Policy change* is the number of political activities respondents indicated they would engage in to change current government policy. *Tax earmarking* is a binary indicator for whether the respondent was exposed to the earmarking prime (treatment) or the general fund prime (control).

This is corroborated by the lack of evidence for any of the underlying mechanisms that could lead earmarking to reduce political participation. To uncover underlying mechanisms, columns 2-4 present the effects of tax earmarking on various motives to engage in political action: information seeking (H3a), discretion reduction (H3b) and policy change (H3c). The tax earmarking treatment did not make an appreciable difference to any of these outcomes; the coefficients are close to zero in magnitude and not statistically significantly different from zero. None of the additionally calculated p-values indicate statistical significance, and the results are robust to covariate adjustment. Curiously, earmarking even appears to have reduced participation for the sake of policy change whereas the opposite was predicted by the theoretical framework.

Additional evidence on the precise mechanisms comes from treatment effect heterogeneity. Table 4.3 presents interaction effects between the tax earmarking treatment and perceptions of government responsiveness beliefs as well as policy benefits. The interaction with perceptions of policy benefits from earmarking is not statistically significant and very small in magnitude (-0.086 and -0.094 on a scale [0, 7]). This suggests that citizens are not more likely to mobilize to change policy in response to earmarked taxation the higher the benefits from the earmarked good (H3d). It also suggests that earmarking does not reduce general participation more the lower the

benefits from the earmarked good (H4b). However, we do observe a positive and statistically significant interaction effect with perceptions of government responsiveness. Tax earmarking *increases* participation more the higher perceptions of government responsiveness. This makes intuitive sense and is one possible explanation for the overall null result: not enough respondents believed that the government would be responsive to their demands, but those that do are more likely to participate. However, this interaction effect goes in a direction opposite from the one that was predicted theory. Hypothesis 4a predicted a negative interaction effect between earmarking and responsiveness on participation, that is, it suggested that tax earmarking *reduces* participation more the higher beliefs about government responsiveness.

Table 4.3: Effect of Tax Earmarking on Political Participation: Mechanisms

	Participation	Participation	Policy change
Intercept	3.987*** (0.267)	3.613*** (0.570)	4.217*** (0.570)
Tax earmarking	-0.581 (0.363)	0.338 (0.803)	0.167 (0.817)
Responsiveness	0.086 (0.088)		
Tax earmarking X Responsiveness	0.199* (0.115)		
Policy benefits		0.155 (0.137)	0.253* (0.137)
Tax earmarking X Policy benefits		-0.086 (0.195)	-0.094 (0.197)
Num. obs.	702	699	741

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Coefficients on *Tax earmarking* are OLS estimates with heteroskedasticity-robust standard errors. *Participation* is the number of political activities respondents indicated they would engage in. *Policy change* is the number of political activities respondents indicated they would engage in to change current government policy. *Tax earmarking* is a binary indicator for whether the respondent was exposed to the earmarking prime (treatment) or the general fund prime (control). *Responsiveness* measures respondents' perception of government responsiveness. *Policy benefits* measures respondents' preference for carrying out government policy through dedicated agencies.

4.5.2 Alternative Mechanisms

These null results are also evident in the fact that there is no evidence for any of the alternative mechanisms that would have predicted either higher or lower participation in response to earmarking. A set of different mechanisms could have been at play,

possibly reversing the predicted relationship between tax earmarking and political participation. To examine this, Table 4.4 presents the effects of tax earmarking on various secondary outcome measures.

Table 4.4: Effect of Tax Earmarking on Political Participation: Alternative Mechanisms

	Responsiveness change	Left out	Policy reversal	Government competence
Tax earmarking	-0.016 (0.064)	-0.035 (0.073)	-0.020 (0.225)	0.042 (0.080)
Control mean	-0.084	3.925	5.178	2.091
Num. obs.	864	867	854	863

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Coefficients on *Tax earmarking* are OLS estimates with heteroskedasticity-robust standard errors. *Responsiveness change* measures the baseline-to-endline change in respondents' perceptions of government responsiveness. *Left out* measures respondents' perceptions of being left out of the political deliberation and policy-making process. *Policy reversal* measures respondents' indicated likelihood that the next government will reverse spending priorities. *Government competence* measures respondents' perception of the government's competence in managing the public finances. *Tax earmarking* is a binary indicator for whether the respondent was exposed to the earmarking prime (treatment) or the general fund prime (control).

First, it could be that citizens' negative update of responsiveness beliefs under general fund taxation is greater than their positive update of beliefs in response to earmarked taxation (H5a). To examine this possibility, I asked the following question both at baseline and at endline and examined the change score: "Do you agree or disagree with the following statement: The government is generally responsive to demands made by people like me." However, there is no evidence that the earmarking treatment affected this change score in any direction since the coefficient magnitude is small (-0.016 with a control group mean of -0.084).

Second, it could be that citizens are more likely to feel left out of the policy deliberation process in response to earmarked taxation than to general fund taxation (H5b). To test this, I examine answers to the following question: "Do you agree or disagree with the following statement: Citizens in Ghana are generally left out of the political deliberation and decision-making process." However, again, there is no evidence that the earmarking treatment affected this perception in any direction; the coefficient magnitude is again close to zero (-0.035 with a control group mean of 3.925).

Third, it could be that citizens are less likely to predict policy reversal of private goods in response to earmarked taxation than to general fund taxation (H5c). To examine this possibility, I look at answers to the question: "On a scale from 0-10, how likely do you think it is that a change in the next government will reverse the current spending policies?" There is again no evidence that the earmarking treatment

affected this prediction in any direction since the coefficient magnitude is close to zero (-0.02 with a control group mean of 5.178).

Finally, it could be that citizens are less likely to evaluate government as fiscally competent in response to earmarked taxation than to general fund taxation (H5d). To test this, I examine answers to the following question: “Do you agree or disagree with the following statement: The government is generally competent in managing the public’s finances.” I find no evidence that the earmarking treatment affected this perception in any direction; the coefficient magnitude is small (0.042 with a control group mean of 2.091).

4.5.3 Explaining Null Results

The above results indicate support for a hypothesis that predicted no difference between treatment groups (H6a). To further examine null results, I pre-registered two additional hypotheses that might explain the absence of a treatment effect of earmarking on political participation.

First, it could be that citizens are less likely to express willingness to pay taxes (exit) in response to general fund taxation than to earmarked taxation (H6b). According to this logic, citizens exit in response to general fund taxation instead of mobilizing, thus producing a null result. To look at this, I examine answers to the following question: “Do you agree or disagree with the following statement: Taxpayers must pay the taxes that they owe to the government, regardless of the quality of public services.” However, I find no evidence that earmarked taxation leads to greater expressed tax morale and thus to less exit; the coefficient is small in magnitude (0.023 with a control group mean of 3.185). Thus, citizens do not exit at higher rates in response to general fund taxation, so this explanation is unlikely to make sense of the null results. However, this piece of evidence hints at the absence of a possible ceiling effect: it could have been the case that it was simply too difficult for the earmarking treatment to move already high average levels of participation in the control group. The fact that respondents do not score lower on the tax morale question in response to the earmarking treatment suggests that they do not try to reduce their tax burden in differential ways. Thus, it does not seem to be the case that respondents were motivated by either treatment to fight taxes politically as is evident from the high levels of agreement with the tax statement in the control group. However, in the absence of a pure control group that did not prime taxation, we cannot rule out these ceiling

effects with certainty. Note also that only 12 percent of respondents scored the highest value of 7 on the main outcome variable which counts instances of participation; 17 percent scored below three, and 39 percent above the mean.

Additionally, it could be that citizens are not more likely to expect other citizens to get involved in politics in response to general fund taxation than to earmarked taxation (H6c). According to this logic, if general fund taxation would really lead to more participation, then heightened beliefs in collective action should accompany that effect. To examine this possibility, I look at answers to the question: “On a scale from 0-10, how likely do you think it is that others in your neighborhood will engage in politics in the future?” I find support for this hypothesis. There is no statistically significant difference between the earmarking treatment and the general fund control group with respect to respondents’ belief in collective action; the coefficient is small in magnitude (0.163 with a control group mean of 5.883). Thus, the earmarking treatment did not depress beliefs in collective action, a process that should have taken place if general fund taxation did in fact increase participation. This explanation for the null result makes intuitive sense with regards to the literature on collective action. For there to be any (even expressed) propensity to participate in political life, citizens have to solve the collective action dilemma that presents when confronted with such coordination problems (Olson, 1971). The only way out of this are the provision of private benefits (Olson, 1971), psychological rewards (Bueno de Mesquita and Shadmehr, 2022) or reduced social costs (Paler, Marshall and Atallah, 2018). General fund taxation seems to offer none of these incentives to overcome collective action dilemmas, at least not in comparison with earmarked taxation.

Table 4.5: Effect of Tax Earmarking on Political Participation: Explanations for Null Results

	Tax morale	Belief in collective action
Tax earmarking	0.023 (0.087)	0.163 (0.189)
Control mean	3.185	5.883
Num. obs.	866	860

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Coefficients on *Tax earmarking* are OLS estimates with heteroskedasticity-robust standard errors. *Tax morale* measures respondents’ willingness to pay taxes. *Belief in collective action* measures respondents’ reported likelihood that others in their neighborhood will engage in politics in the future. *Tax earmarking* is a binary indicator for whether the respondent was exposed to the earmarking prime (treatment) or the general fund prime (control).

4.5.4 Heterogeneity

As hypothesized in the pre-analysis plan, treatment effects might vary by co-ethnicity with the main ethnic group, co-partisanship with the incumbent presidential party, or respondent gender. Table 4.6 presents the interaction effects of the tax earmarking treatment with the respective covariate of interest. Here, we restrict the analysis to the main outcome variable of interest, general political participation.

First, I find no evidence that the earmarking treatment effect varies by co-ethnicity with the main ethnic group. The interaction effect is small in magnitude (0.393 with a control group mean of 4.459) and insignificant. Second, I find equally no evidence for heterogeneity according to support for the incumbent presidential party; again, the coefficient of interest is small in magnitude (0.420 with a control group mean of 5.167) and statistically insignificant. Finally, I do find evidence that the treatment might have affected respondents differently depending on their gender. The interaction effect with a binary variable for female is negative and statistically significant. Thus, among female respondents, earmarking appears to have reduced participation (by 0.642 points with a control group mean of 4.473) as predicted in the main hypothesis of the theoretical framework. Investigating this further and unpacking this effect heterogeneity to theorize why earmarking affects men and women differently is a fruitful task for future research.

4.6 Discussion

In this chapter, I have provided a theoretical clarification of the relationship between the tax-benefit linkage and political participation. While most of the existing literature thinks that earmarking taxes has positive consequences for the demand for accountability, I specify conditions which would make us reach conclusions opposite to the current way of thinking about the political consequences of earmarking. I present preliminary evidence that tax earmarking may make no difference to citizens' propensity to participate in political life. In a survey experiment conducted in Ghana, presenting respondents with a vignette that primes tax earmarking does not affect the number of political activities they would engage in. The effect magnitude here is close to zero and the hypothesized mechanisms for any effect do not seem to be at play. Two explanations for this null result emerged from the results: citizens may have lacked belief in (i) general government responsiveness to citizen efforts at influencing political life, and (ii) the collective action capacity of society as a whole

Table 4.6: Effect of Tax Earmarking on Political Participation: Heterogeneity

	Participation	Participation	Participation
Intercept	4.459*** (0.167)	5.167*** (0.282)	4.473*** (0.146)
Tax earmarking	-0.209 (0.226)	-0.128 (0.385)	0.278 (0.195)
Main ethnicity	-0.437** (0.218)		
Tax earmarking X Main ethnicity	0.393 (0.304)		
Incumbent supporter		-0.550 (0.370)	
Tax earmarking X Incumbent supporter		0.420 (0.505)	
Female			-0.548** (0.214)
Tax earmarking X Female			-0.642** (0.296)
Num. obs.	697	196	702

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Coefficients on *Tax earmarking* are OLS estimates with heteroskedasticity-robust standard errors. *Participation* is the number of political activities respondents indicated they would engage in. *Tax earmarking* is a binary indicator for whether the respondent was exposed to the earmarking prime (treatment) or the general fund prime (control). *Main ethnicity* measures respondents' co-ethnicity with the main ethnic group. *Incumbent supporter* measures respondents' co-partisanship with the incumbent presidential party. *Female* is a binary indicator for whether the respondent identifies as female.

to overcome coordination problems that surround political participation. It is also evident that earmarking did in fact reduce participation among female respondents.

However, it is important to bear in mind several limitations of this chapter. First, several theoretical ideas were deliberately not investigated as lying outside the scope of the project. For example, it was assumed that citizens possess sufficient information to determine alignment of their preferences with government policy, but this was not explicitly tested. Similarly, it was assumed but not explicitly tested that citizens actually believe a promise of the government to earmark tax revenue. Second, the survey experiment did not directly manipulate the actual extent of tax earmarking and thus citizens' perceptions thereof, nor did it measure actual participation in political life. All that was done was to present respondents with a vignette as a prime on tax earmarking and then measure *stated* preferences regarding participation. One explanation for the null results might simply be that there were no real-world costs and benefits involved in the experiment so that the experimental environment was

simply too artificial. Finally, the survey was only conducted in a single country, Ghana, from which it may be difficult to extrapolate results to other contexts which have not had the same kind of history of contestation surrounding tax policy and especially tax earmarking. The results might be different in, for example, Mexico, where research has shown that earmarked taxes are preferred to general fund taxation (Flores-Macías, 2018).

Regardless of these limitations, the results presented in this chapter run counter to expectations in the literature on taxation and accountability, as well as to predictions obtained from a simple decision-theoretic framework. On the one hand, several scholars have argued that tax earmarking might lead to greater citizen demands for government accountability due to improved transparency of the tax price of public goods. On the other hand, a simple decision-theoretic framework based on standard political economy assumptions shows that one can obtain opposite predictions, namely that earmarking reduces participation. However, I find no preliminary evidence for either of these perspectives. Earmarking does not seem to engage the citizenry to a greater or lesser extent than general fund taxation. This suggests that we need to revise our (positive or negative) view of tax earmarking and its implications for political participation. Given the clear limitations of the empirical test employed in this chapter, it also suggests the need for more empirically robust research that tests the intuitions and observable implications derived in the theoretical framework. The preliminary evidence deployed in this chapter is certainly not the final word on the subject.

The main implication of the results so far is that governments can continue to expand general fund taxation without having to fear political costs in the form of societal mobilization. If this is correct, governments can harmonize their desire to maintain or increase spending discretion with anticipated political responses. In particular, governments do not need to (i) have accurate information about likely contestation, (ii) preempt mobilization by strategically releasing tax-and-spend information, (iii) care about political survival when deciding whether to earmark or not, and (iv) tie their political survival to societal mobilization. Overall, this might help explain why general fund taxation is the more common form of the tax-benefit linkage today.

Chapter 5

Conclusion

This thesis has intended to shed light on questions of state-society relations with experimental evidence from a range of developing countries on the effects of state capacity building on citizens' political behavior. Chapter 2 asked whether delegating tax collection to local elites such as city chiefs in the Democratic Republic of the Congo has adverse consequences for bottom-up accountability pressure by citizens. In this chapter, I documented that delegating tax collection to local elites does not impede citizens' ability or willingness to hold both chiefs and the state accountable for policy outcomes. Chapter 3 investigated in India and the United Kingdom whether promising to fight corruption as a political candidate generates electoral rewards by voters. In this chapter, I showed that promising to fight corruption as a central policy platform is electorally punished by voters compared to promising to improve the economy. Chapter 4 examined the accountability consequences of earmarking tax revenue for particular public expenditures in Ghana. In this chapter, I tentatively showed that citizens might engage in political life in response to taxation to the same degree, regardless of whether tax revenue is earmarked or not.

By examining the micro-behavioral foundations of big theoretical claims in the literature on state-society relations, this thesis documents partially sizeable and significant political responses to the state's effort at increasing its capacity, either in the form of delegating taxes or fighting corruption. When citizens play an active role in shaping state-society relations, as this thesis suggests they oftentimes do, increasing state capacity may not necessarily lead to autocratic forms of governance. Instead, the state's attempt at increasing its reach may spark bottom-up accountability pressures and citizen mobilization that aim at reigning in the state's power. This is exactly in line with the central argument put forward by Acemoglu and Robinson (2019): society's active involvement in political life constraints the power of the state and creates the virtuous "shackled Leviathan" who can resolve conflicts, control feuds

and provide public services while society is empowered to take control over the state and transform prevailing norms of contestation. At other times, for example when earmarking taxes, increasing capacity may not significantly affect citizen participation. Greater nuance in the precise sequencing of political development is therefore needed, and this thesis intends to take a first step towards this goal.

This thesis documents political responses by civil society to increases in state capacity in both developed and developing countries. What does this imply? While local conditions can differ by context—in terms of levels of state capacity, the role of elites, initial levels of political participation—the broader lessons for how taxation and corruption structure political engagement are plausibly general. State capacity runs along a continuum and remains dynamic; it can increase or decrease in any given context with any given level of existing capacity. However, as this thesis shows, the political behavior responses are broadly similar. Nevertheless, certain scope conditions constrain the applicability of the findings presented here: a set of meaningful democratic practices, a modicum of state capacity (for changes in the way the state taxes to take place and for fighting corruption to be a credible promise), and a collection of elites working closely with the state, not against it.

Each chapter leaves open a number of questions to be addressed by future research. Chapter 2 showed that delegating tax collection to local non-state actors can raise revenue without an associated accountability penalty and the impediment of the formation of a citizen-state fiscal contract. However, it is an open question to what extent this is true for customary chiefs as well as other local elites that were not incorporated into the colonial government. In addition, we do not know what would happen in contexts in which not just tax collection but also spending authority was delegated, and in which ethnicity was a more salient cleavage. Chapter 3 demonstrated significant possible electoral costs to political campaigns that center on fighting corruption as opposed to improving the economy. However, to what extent a candidate would actually lose a real-world election cannot be inferred from the results presented here. Conjoint experimental designs do not allow for such an interpretation, especially when a uniform distribution of candidate characteristics is assumed. In addition, we do not know what would happen in a more realistic scenario where policy is multidimensional and candidates could campaign on multiple issues simultaneously. Chapter 4 preliminarily exposed the lack of sizeable positive accountability consequences of earmarking taxes for particular spending purposes. However, from the results presented here we cannot infer whether this is true of all earmarked policy items or just for public health as in the experiment. In addition, from the

stated preferences in the survey we do not know if citizens would mobilize in the real world in response to general fund taxation, and how they would overcome problems of collective action. Closing these knowledge gaps is a fruitful task for future research.

A main open question of this thesis as a whole is its implications for the literature on state building (as opposed to state formation). In particular, the role of elites as well as elite-mass relations could be examined in closer detail. With respect to the role of elites, Wang (2021) argues for a blurring of state-society boundaries to incorporate elites into the analytical framework of state building. In his model, elite social terrains—different degrees of embeddedness in society—determine their efforts in state building. For example, city chiefs in the DRC could be depicted either in a bowtie or ring network, to use Wang’s terminology. Elite preferences are then tied to particular constituencies along salient cleavages, in the context of many sub-Saharan African states possibly ethnicity. A question also remains about elite incorporation into the state. Garfias and Sellars (2023) argue that if elites are incorporated, their benefits from state survival increase, and they thus might be willing to accept higher taxation in return. With respect to elite-mass relations, one unaddressed issue is the role of the threat of commoner rebellion. Garfias and Sellars (2023) show that when the threat of commoner rebellion increases (in our case through increases in societal mobilization), elites have greater bargaining power vis-à-vis the state but are also cross-pressured by the desire to fend off rebellion. Thus, the threat of rebellion has ambiguous effects on state building; it then depends on other factors such as the preference alignment of elites and the state. Engaging with this line of reasoning is a major task for future research.

Appendices

Appendix A

Delegating Tax Collection Does Not Adversely Affect Demand for Accountability: Evidence from an Experiment in the D.R. Congo

A.1 Additional Information on Experimental Design

A.1.1 Intervention and Data Collection Activities

Treatments were randomly assigned at the neighborhood level. In this paper, we use the subset of Gov and Chief neighborhoods in Kananga that had a collective action opportunity ($J = 106$ in total including the pure control group), i.e., neighborhoods in which citizens received fliers for meeting invitations as well as the accompanying meeting request forms. Within those neighborhoods, $N = 2815$ households received a flier; for these households, we were able to verify correct audit form submission or absence of submission. Assignment was stratified on a combination of geographic location as well as tax compliance in the previous tax campaign. We report the timeline of the main intervention and data collection activities in Table A.2. Details on the full intervention as well as on the lottery ticket distribution for the anti-poverty program are reported in Balán et al. (2022) and Bergeron et al. (2022).

Table A.1: Treatment Allocation in the 2018 Property Tax Campaign

		State	Chief	Control
<i>Tax Campaign Sample</i>	Neighborhoods	110	111	5
	Property owners	14,489	14,383	797
<i>Collective Action Sub-Sample</i>	Neighborhoods	50	51	5
	Flier recipients	1,314	1,317	169

Table A.2: Timing of all Activities and Data Collection

Activity	Actor	Timing	N	J
Tax campaign				
<i>Taxation</i>				
Property registration	Collectors	May–Dec 2018	13,996	106
Tax visits	Collectors	May–Dec 2018	13,996	106
<i>Evaluation</i>				
Baseline survey	Enumerators	Jul–Dec 2017	2,268	106
Midline survey	Enumerators	Jun 2018–Feb 2019	10,928	106
Endline survey I	Enumerators	Mar–Sep 2019	1,626	106
Collective action opportunity				
<i>Program</i>				
Flier and audit form distribution	SOCICO	Jun–Oct 2019	2,815	106
<i>Evaluation</i>				
Audit form collection and validation	Enumerators	Oct 2019	2,815	106
Cash transfer program				
<i>Program</i>				
Program ticket distribution	Chiefs	Jun–Oct 2019	2,799	106
Lottery	Chiefs & DIVAS	Jun–Oct 2019	106	106
Cash transfer distribution	Chiefs	Jun–Oct 2019	530	106
<i>Evaluation</i>				
Endline survey II	Enumerators	Jun–Dec 2019	3,037	106

A.1.2 Audit Request Form

Figure A.1: Audit Request Form

Do you want an audit and verification meeting?
As part of this program, you and other people in your avenue can **request an audit and verification meeting** organized by a civil society organization in Kananga. This is an opportunity for you as a [citizen/taxpayer] to learn more about this program and whether it was implemented properly and fairly. The meeting can focus on the actions taken by the **Division of Social Affairs**, by your **avenue chief**, or **both** in this development program.

IMPORTANT : The civil society organization will only organize a meeting for your avenue if many residents request one.

- To **request an audit meeting of [Actor1]**, submit the [COLOR] form to the [COLOR] drop box located at [ADDRESS1].
- To **request an audit meeting of [Actor2]**, submit the [COLOR] form to the [COLOR] drop box located at [ADDRESS2].

To request meetings involving **both** actors, submit both forms to the correct boxes. Everything you write will be kept confidential from the concerned parties. All forms must be submitted by [date].

The avenues that submit the most requests (as a share of all households) will get top priority to receive an audit meeting. Your action is important!

**REQUEST MEETING
of the [DIVAS/Chef].**

To request a meeting of the [DIVAS/Chef], please **deposit this form into the locked box at :**

[LOCATION].

The box will have show this colored stamp:

[COLOR STAMP]

Request of the compound: [Code]

**REQUEST MEETING
of the [DIVAS/Chef].**

To request a meeting of the [DIVAS/Chef], please **deposit this form into the locked box at :**

[LOCATION].

The box will have show this colored stamp:

[COLOR STAMP]

Request of the compound: [Code]

A.1.3 Design Diagnostics: Balance and Interference

In Tables A.3 and A.4, we report results of a series of balance tests on important pre-treatment covariates. In Table A.3, Panels A-C test balance for randomization of the tax collection treatment in all neighborhoods, and panels D-F test balance for the randomization of flier receipt. In Table A.4, we repeat this analysis for flier recipients only and examine balance for the tax collection treatment. In particular, we report a series of regressions of the main covariate of interest on the treatment indicator. These covariates include characteristics of property owners, properties and neighborhoods. Table A.5 reports on the extent of interference by examining outcomes from the endline survey.

Table A.3: Balance Tests for Covariates

	N	J	$\bar{x}_C$	$\hat{\beta}$	SE	p
I. Gov vs Chief Tax Collection						
Panel A: Property owner characteristics						
Age	1203	101	50.59	-0.55	1.17	0.64
Female	1203	101	0.35	-0.04	0.03	0.18
Has electricity	1203	101	0.15	-0.02	0.02	0.35
Erosion	1202	101	0.25	0.05	0.06	0.39
Fence quality	1202	101	1.47	0.01	0.08	0.92
Years of education	1201	101	10.87	-0.51	0.39	0.19
Log HH monthly income	1193	101	10.60	0.05	0.25	0.84
Possessions (wealth)	1203	101	1.16	-0.03	0.11	0.81
Employed	7372	100	0.73	0.00	0.02	0.93
Salaried	7375	101	0.26	-0.01	0.01	0.46
Works for government	7375	101	0.17	-0.01	0.01	0.55
Majority tribe	6843	101	0.77	0.03	0.04	0.46
Panel B: Property characteristics						
Walls good cond.	10801	101	1.75	-0.03	0.05	0.60
Roof good cond.	10420	101	0.44	-0.02	0.04	0.60
Dist. city center	13064	101	3.01	0.03	0.27	0.93
Dist. state buildings	13064	101	0.77	0.16	0.12	0.19
Dist. hospitals	13064	101	1.18	0.02	0.16	0.91
Dist. public schools	13064	101	0.37	0.04	0.05	0.35
Panel C: Neighborhood characteristics						
P.c. property tax revenue 2016		101	104.69	11.57	39.58	0.77
Avg. participation index 2017		101	-0.02	-0.04	0.07	0.49
II. Flier Receipt						
Panel A: Property owner characteristics						
Age	1203	101	50.39	-0.19	1.11	0.86
Female	1203	101	0.32	0.06	0.03	0.08
Has electricity	1203	101	0.14	-0.01	0.03	0.64
Erosion	1202	101	0.28	0.01	0.03	0.64
Fence quality	1202	101	1.49	-0.03	0.05	0.58
Years of education	1201	101	10.62	0.21	0.32	0.52
Log HH monthly income	1193	101	10.66	-0.04	0.19	0.84
Possessions (wealth)	1203	101	1.16	-0.03	0.09	0.78
Employed	7372	100	0.73	0.00	0.01	0.74
Salaried	7375	101	0.26	0.01	0.01	0.47
Works for government	7375	101	0.16	0.01	0.01	0.30
Majority tribe	6843	101	0.79	0.00	0.01	0.88
Panel B: Property characteristics						
Walls good cond.	10801	101	1.74	-0.01	0.02	0.58
Roof good cond.	10420	101	0.43	0.00	0.01	0.73
Dist. city center	13064	101	3.04	-0.01	0.01	0.70
Dist. state buildings	13064	101	0.86	0.00	0.01	0.84
Dist. hospitals	13064	101	1.21	-0.01	0.01	0.15
Dist. public schools	13064	101	0.39	0.00	0.00	0.42
Panel C: Neighborhood characteristics						
P.c. property tax revenue 2016		101	113.01	1.74	1.39	0.22
Avg. participation index 2017		101	-0.06	0.00	0.00	0.17

Note: We report here a series of individual regressions of the main covariate of interest on the treatment indicator. These covariates include characteristics of property owners, properties and neighborhoods. Subpanels A-C test balance for (i) the randomization of the tax collection treatment (Chief v Gov) and (ii) for randomization of flier receipt. For all variables except the neighborhood-level variables in Panel C, we use individual receipt of a flier within collective action neighborhoods as the predictor. In Panel C, we use the number of fliers per polygon as the predictor since this is a polygon-level outcome. As usual, these regressions include cluster-robust standard errors and randomization stratum fixed effects. In Panel II.C, we use WLS to regress cluster-level averages on treatment indicators, with weights proportional to cluster size. Note that for the midline variables in Panels I.B and II.B, 44-48% of the 13,267 observations are missing.

Table A.4: Balance Tests For Covariates Among Flier Recipients

	N	J	$\bar{x}_C$	$\hat{\beta}$	SE	p
Panel A: Property owner characteristics						
Age	256	94	51.20	-2.10	2.14	0.33
Female	256	94	0.36	0.03	0.06	0.58
Has electricity	256	94	0.15	-0.05	0.04	0.19
Erosion	256	94	0.23	0.10	0.08	0.19
Fence quality	256	94	1.47	-0.06	0.10	0.53
Years of education	255	94	10.94	-0.57	0.55	0.31
Log HH monthly income	252	94	10.48	0.14	0.34	0.68
Possessions (wealth)	256	94	1.15	-0.09	0.18	0.61
Employed	1460	98	0.74	0.00	0.03	0.92
Salaried	1462	99	0.27	-0.01	0.02	0.66
Works for government	1462	99	0.17	-0.01	0.02	0.73
Majority tribe	1341	99	0.77	0.02	0.04	0.62
Panel B: Property characteristics						
Walls good cond.	2239	101	1.73	0.01	0.05	0.89
Roof good cond.	2161	101	0.43	-0.01	0.04	0.83
Dist. city center	2611	100	3.00	0.04	0.28	0.89
Dist. state buildings	2611	100	0.77	0.15	0.12	0.21
Dist. hospitals	2611	100	1.17	0.03	0.17	0.88
Dist. public schools	2611	100	0.37	0.04	0.05	0.35

Note: We report here a series of individual regressions of the main covariate of interest on the treatment indicator for polygons assigned to the chief collection treatment. We restrict the sample here to flier recipients, i.e. those households that had a collective action opportunity, and assess balance across tax collection treatments within this subset. Covariates include characteristics of property owners and properties. Note that we are unable to link household-level flier receipt to the households that were part in the 2016 tax campaign, and so we don't assess balance on those neighborhood-level characteristics. Panels A and B test balance for the randomization of the tax collection treatment (Chief v Gov). As usual, these regressions include cluster-robust standard errors and randomization stratum fixed effects. Note that for the midline variables, 44-48% of the 13,267 observations are missing.

Table A.5: Extent of Interference in Treatments

	Heard about chief taxation		Heard about government taxation	
	(1)	(2)	(3)	(4)
Chief tax collection	0.329*** (0.054)	0.280*** (0.058)	-0.180*** (0.055)	-0.121** (0.058)
Control mean	0.259	0.259	0.819	0.819
Observations	577	576	577	576
Clusters	86	86	86	86
Stratum FE	Yes	Yes	Yes	Yes
Controls	No	Set I	No	Set I
Sample	C + L	C + L	C + L	C + L

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Coefficients on *Chief tax collection* are OLS estimates with robust standard errors clustered at the polygon level, the unit of randomization and of analysis. Outcomes come from the Endline I survey. *Heard about chief taxation* is an indicator that is 1 if the respondent said they had heard about the chief working as tax collector for the 2018 campaign anywhere in Kananga. *Heard about government taxation* is an indicator that is 1 if the respondent said they had heard about the government working as tax collector for the 2018 campaign anywhere in Kananga. *Chief tax collection* is a binary indicator for whether the polygon experienced tax collection by the chief (1) or by the provincial government (0; control). Here we compare polygons that had a collective action opportunity ($j = 101$) and where either tax collection was done exclusively by the chief ($j = 51$) or by the provincial government ($j = 50$). We restrict the sample here further to exclude outliers on the capacity questions as well as villas. *Stratum FE* refer to the stratum used for randomization which is a combination of geographic location of the neighborhood as well as previous tax compliance. *Set I* controls include house type, wave of collective action campaign, distances to letter boxes and polygon-level averages of participation and per capita tax revenue in 2016.

A.2 Appendices for Main Results

A.2.1 Robustness Checks

We show that our main results are robust to a wide variety of alternative specifications and additional tests. In Table A.6, we add three additional sets of control variables to each of the primary outcomes we presented in Table 2.1. These include a range of prognostic pre-treatment covariates at the level of the polygon, house, and individual citizen. In Table A.7, we repeat the main analysis using different randomization strata at a more fine-grained level which stratifies on the geographic location of neighborhoods. These strata could alternatively have been used due to the nature of the cross-randomization with other factors of the experimental design fully reported in Bergeron et al. (2022).

Table A.6: Effects of Chief Tax Collection on Demand for Citizen Meetings: More Controls

	Gov meeting			Chief meeting			Either			Both		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Chief tax collection	-0.001 (0.030)	0.001 (0.039)	-0.004 (0.037)	0.002 (0.031)	0.009 (0.041)	0.005 (0.041)	0.005 (0.032)	0.007 (0.041)	0.003 (0.041)	-0.004 (0.030)	0.002 (0.038)	-0.002 (0.038)
Control mean	0.169	0.169	0.169	0.171	0.171	0.171	0.180	0.180	0.180	0.161	0.161	0.161
Observations	2160	1458	1397	2160	1458	1397	2160	1458	1397	2160	1458	1397
Clusters	101	98	98	101	98	98	101	98	98	101	98	98
Stratum FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Set II	Set III	Set IV	Set II	Set III	Set IV	Set II	Set III	Set IV	Set II	Set III	Set IV

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Coefficients on *Chief tax collection* are OLS estimates with robust standard errors clustered at the polygon level, the unit of randomization and of analysis. *Chief Meeting* is the proportion of citizens requesting a meeting with the chief in a given polygon. *Gov Meeting* is the proportion of citizens requesting a meeting with the provincial government in a given polygon. *Either* is the proportion of citizens in a given polygon requesting either a meeting with the chief or the provincial government. *Both* is the proportion of citizens in a given polygon requesting both a meeting with the chief and with the provincial government. *Chief tax collection* is a binary indicator for whether the polygon experienced tax collection by the chief (1) or by the provincial government (0; control). Here we compare polygons that had a collective action opportunity ($j = 101$) and where either tax collection was done exclusively by the chief ($j = 51$) or by the provincial government ($j = 50$). *Stratum FE* refer to the stratum used for randomization which is a combination of geographic location of the neighborhood as well as previous tax compliance. *Set II* controls include Set I controls plus wall and roof material of the compound. *Set III* controls include Set I controls plus binary indicators for the property owner being employed, salaried and working for the government. *Set IV* controls include Set I controls plus age and gender of the property owner.

Table A.7: Effects of Chief Tax Collection on Demand for Citizen Meetings: Alternative Strata

	Gov meeting		Chief meeting		Either		Both	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Chief tax collection	-0.007 (0.028)	0.003 (0.026)	-0.003 (0.029)	0.007 (0.027)	-0.001 (0.029)	0.010 (0.027)	-0.009 (0.028)	0.001 (0.026)
Control mean	0.169	0.169	0.171	0.171	0.180	0.180	0.161	0.161
Observations	2631	2629	2631	2629	2631	2629	2631	2629
Clusters	101	101	101	101	101	101	101	101
Stratum FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Set I	No	Set I	No	Set I	No	Set I

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Coefficients on *Chief tax collection* are OLS estimates with robust standard errors clustered at the polygon level, the unit of randomization and of analysis. *Chief Meeting* is the proportion of citizens requesting a meeting with the chief in a given polygon. *Gov Meeting* is the proportion of citizens requesting a meeting with the provincial government in a given polygon. *Either* is the proportion of citizens in a given polygon requesting either a meeting with the chief or the provincial government. *Both* is the proportion of citizens in a given polygon requesting both a meeting with the chief and with the provincial government. *Chief tax collection* is a binary indicator for whether the polygon experienced tax collection by the chief (1) or by the provincial government (0; control). Here we compare polygons that had a collective action opportunity ($j = 101$) and where either tax collection was done exclusively by the chief ($j = 51$) or by the provincial government ($j = 50$). *Stratum FE* refers to the alternative randomization stratum of geographic location of the neighborhood. *Set I* controls include house type, wave of collective action campaign, distances to letter boxes and polygon-level averages of participation and per capita tax revenue in the 2016 campaign.

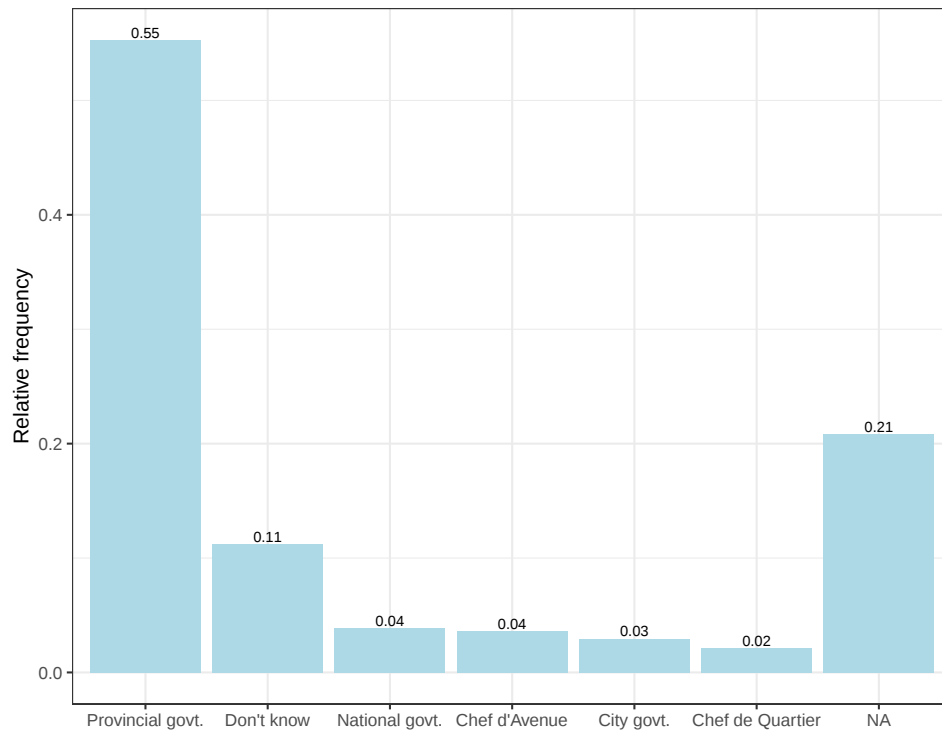
A.2.2 Additional Analyses

Table A.8: Effect of Chief Tax Collection on Civic Participation

	Community meeting		CSO meeting		Political party meeting	
	(1)	(2)	(3)	(4)	(5)	(6)
Chief tax collection	−0.032 (0.039)	−0.031 (0.036)	−0.021 (0.038)	−0.022 (0.036)	−0.015 (0.034)	−0.001 (0.035)
Control mean	0.442	0.442	0.346	0.346	0.304	0.304
Observations	963	962	963	962	963	962
Clusters	101	101	101	101	101	101
Stratum FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Set I	No	Set I	No	Set I
Sample	C + L	C + L	C + L	C + L	C + L	C + L

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Coefficients on *Chief tax collection* are OLS estimates with robust standard errors clustered at the polygon level, the unit of randomization and of analysis. Outcomes come from the Endline I survey. *Community meeting* is an indicator that is 1 if the respondent said they had attended a community meeting in the past year. *CSO meeting* is an indicator that is 1 if the respondent said they had attended meeting by a civic association in the past year. *Political party meeting* is an indicator that is 1 if the respondent said they had attended a meeting by a political party in the past year. *Chief tax collection* is a binary indicator for whether the polygon experienced tax collection by the chief (1) or by the provincial government (0; control). Here we compare polygons that had a collective action opportunity ($j = 101$) and where either tax collection was done exclusively by the chief ($j = 51$) or by the provincial government ($j = 50$). We restrict the sample here further to exclude outliers on the capacity questions as well as villas. *Stratum FE* refer to the stratum used for randomization which is a combination of geographic location of the neighborhood as well as previous tax compliance. *Set I* controls include house type, wave of collective action campaign, distances to letter boxes and polygon-level averages of participation and per capita tax revenue in 2016.

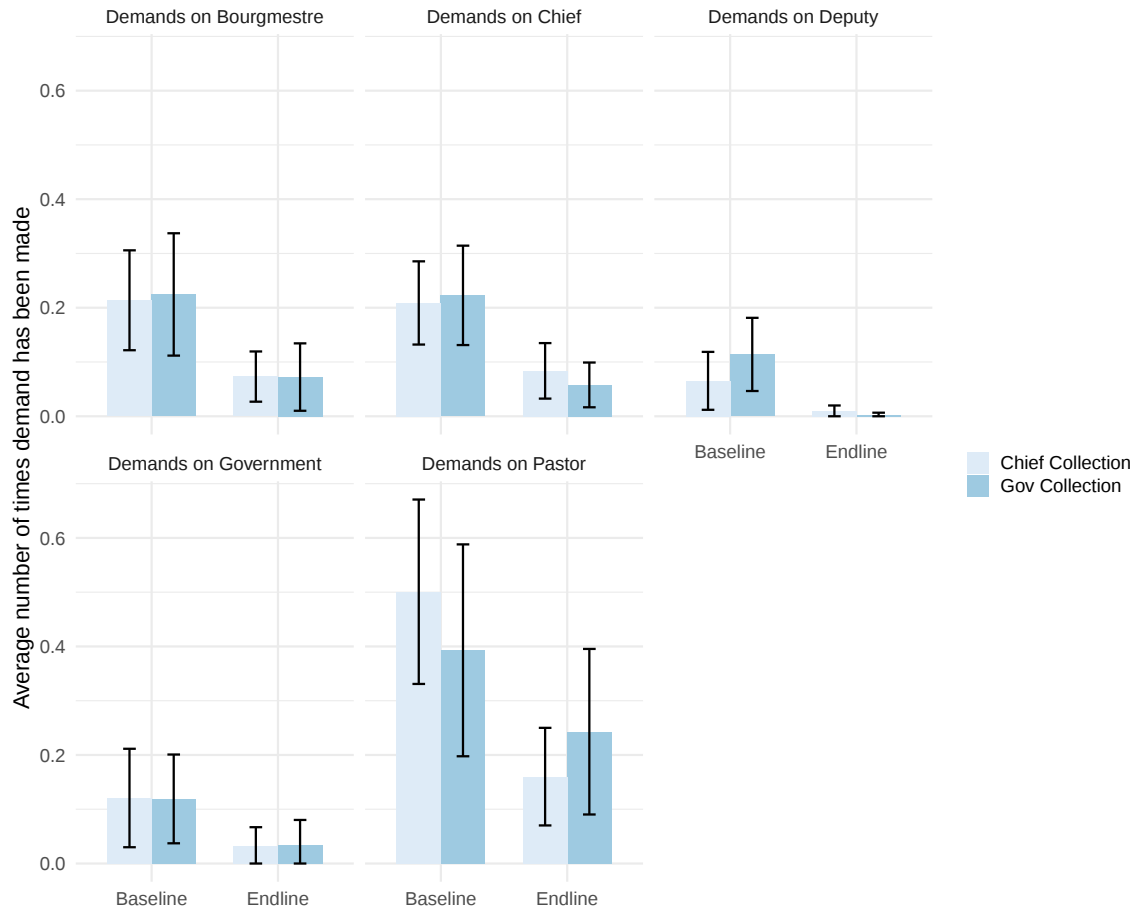
Figure A.2: Recall of Tax Payments Remitted to Which Actor



Notes: We plot the relative frequency of answers to the endline question: “Let’s discuss the property tax. To whom is the tax owed?” We collected data from 1,217 respondents at this endline survey.

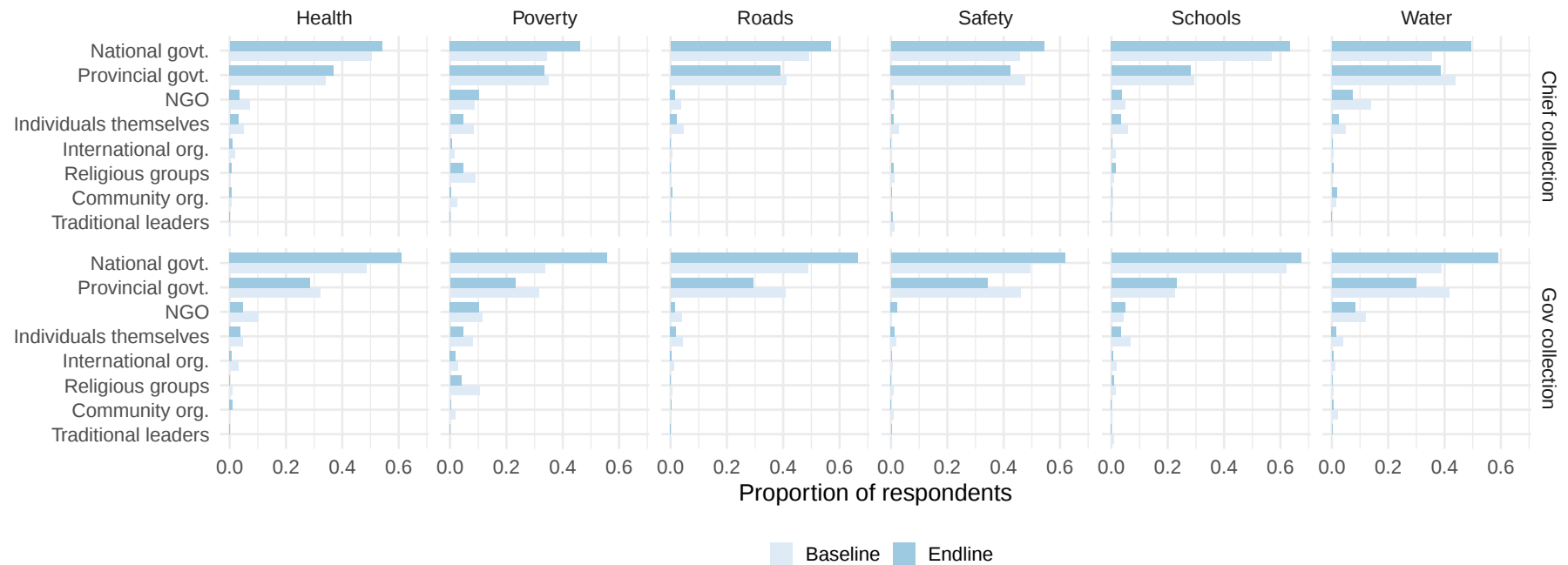
A.2.3 Mechanisms

Figure A.3: Number of Demands on Local Actors



Notes: We plot the average number of times a demand has been made on the respective actor. We measured this at baseline and endline with the question: “In the past 12 months, how many times has a member of your household gone to each of the following people or places to discuss a problem or make a demand?” We collected data from 2,183 respondents at baseline and from 1,217 respondents at endline. Confidence intervals are based on cluster-robust standard errors.

Figure A.4: Responsibility for Local Public Goods Provision



Notes: We plot the relative frequency of answers to the baseline and endline questions: “I am going to list some services/infrastructure many communities have. Tell me who you think should be primarily responsible for providing each one in our community. This does not need to be the current provider of these services/infrastructure.” We collected data from 2,183 respondents at baseline and from 1,552 respondents at endline.

A.3 Appendices for Alternative Explanations

A.3.1 Meetings and Genuine Accountability

We document that the citizen meetings which households in Kananga could request were seen as a meaningful opportunity to exert bottom-up pressures and generate genuine accountability for local public goods provision in the context of an anti-poverty program. First, in Table A.9, we report results of quantitative analysis of text obtained from oral comments made by citizens during the meetings which we summarize as seeded-LDA topic models and show the 10 most common words for 5 different topics of comments. We also show that form submission was a costly activity that citizens took seriously as ways to request actual meetings. We document this by examining how strongly form submission predicts citizen meeting attendance, regressions we perform in Table A.10 for each actor separately. Second, we assemble additional evidence from quantitative text analyses on underlying mechanisms of our main results. We do this by analyzing additional text documents obtained at the actual citizen meetings. In Figure A.6, we report on the main post-processed topics and comments that citizens made during the meetings and examine their frequency both in the aggregate as well as differentiated by treatment group. Figure A.7 takes the actual original oral comments as text corpus and documents the 25 most common words spoken during the meetings, again both in the aggregate as well as differentiated by treatment group.

Table A.9: LDA Topic Model for Comments in Citizen Meetings

Topic 1	Topic 2	Topic 3	Topic 4	Topic 5
argent	injustice	comment	recu	recu
aide	chef	reunion	pourquoi	rien
aussi	aide	pauvrete	argent	connais
tickets	pourquoi	contre	projet	oui
avoir	gens	venu	prix	lotterie
paurquoi	distribution	lutte	autres	argent
tout	corruption	socico	chose	tickets
discrimination	tous	combattre	quelque	ticket
monde	gagne	but	comment	vu
donne	bon	bien	aussi	corruption

Note:

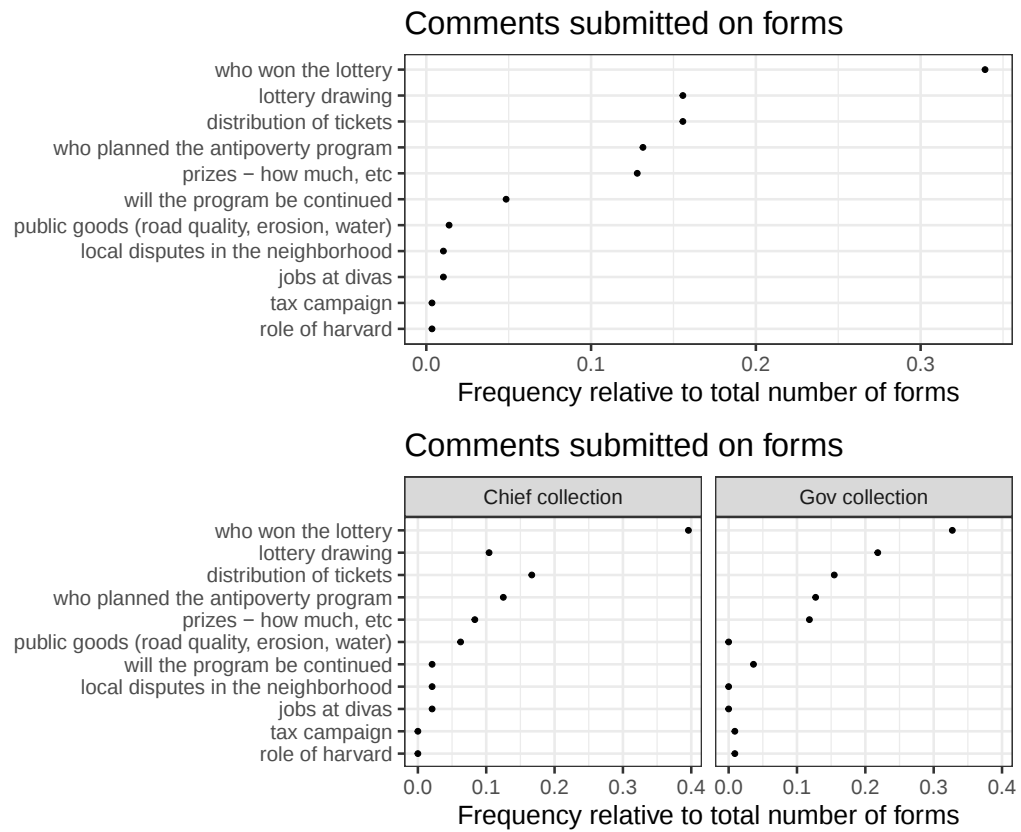
We post-process the text data from the comments made in citizen meetings into a corpus of French language comments, clean strings and remove common stopwords in French. We plot this for all citizens who attended a citizen meeting and made a comment. Based on this corpus, we perform semi-supervised Latent Dirichlet allocation (seeded-LDA) and show here the 10 most common words of 5 topics of comments in citizen meetings.

Table A.10: Meeting Attendance Rates by Form Submission

	Citizen meeting attendance			
	(1)	(2)	(3)	(4)
Submitted form gov	0.090*** (0.014)	0.086*** (0.014)		
Submitted form chief			0.090*** (0.014)	0.086*** (0.014)
Control mean	0.010	0.010	0.010	0.010
Observations	2631	2629	2631	2629
Stratum FE	Yes	Yes	Yes	Yes
Controls	No	Set I	No	Set I

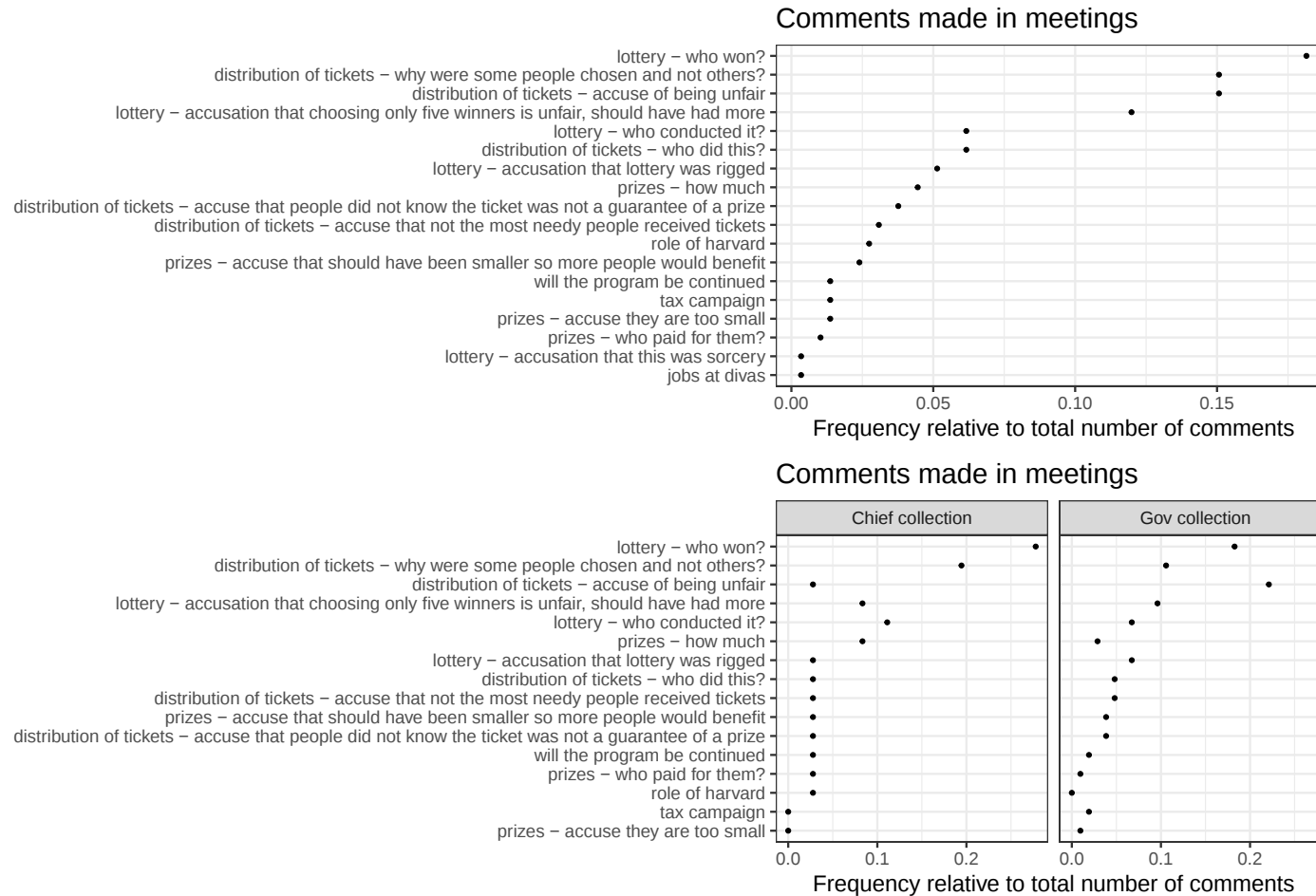
*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Coefficients on *Submitted form* are OLS estimates with HC1 standard errors for individual form submission, the unit of analysis. *Meeting attendance* is 1 if a citizen attended the meeting and 0 otherwise. *Submitted form gov* is a binary indicator for whether the citizen submitted a government meeting request form (1). *Submitted form chief* is a binary indicator for whether the citizen submitted a chief meeting request form (1). *Stratum FE* refer to the stratum used for randomization which is a combination of geographic location of the neighborhood as well as previous tax compliance. The sample here is all citizens that had an opportunity to submit meeting request forms in govt. or chief polygons. *Set I* controls include house type, wave of collective action campaign, distances to letter boxes and polygon-level averages of participation and per capita tax revenue 2016.

Figure A.5: Meeting Request Forms Comments



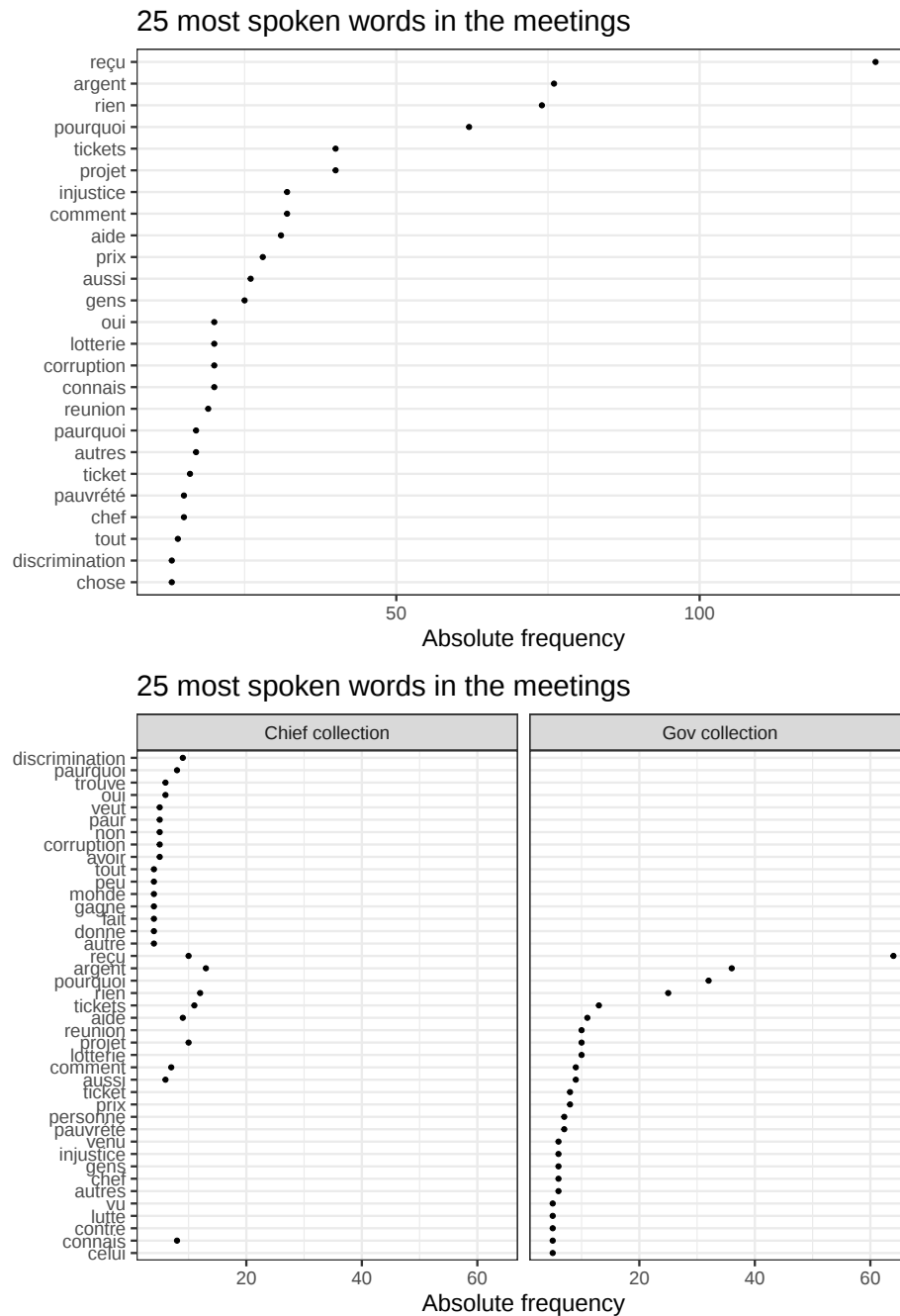
Notes: In the top panel, we plot the frequency of hand-coded comment categories that citizens made. Out of 1,658 forms submitted in all neighborhoods, 289 made additional comments on the form which we plot here as a fraction of the total number of forms with valid comments. In the bottom panel, we plot this same distribution separately for citizens that live in *Gov* or *Chief* neighborhoods as a fraction of total number of forms with valid comments from citizens in *Gov* ($N = 110$) and *Chief* ($N = 48$) neighborhoods, respectively.

Figure A.6: Most Common Topics in Citizen Meetings



Notes: In the top panel, we plot the frequency of hand-coded comment categories that citizens made in citizen meetings. Here, we plot the frequency of these comment categories relative to the total number of 140 comments made during the meeting. In the bottom panel, we plot this same distribution separately for citizens that live in *Gov* or *Chief* neighborhoods as a fraction of total number of comments during meetings from citizens in *Gov* ($N = 104$) and *Chief* ($N = 36$) neighborhoods, respectively.

Figure A.7: Most Common Words in Citizen Meetings



Notes: In the top panel, we plot the frequency of the most common words spoken during citizen meetings. Here, we plot the frequency of words relative to the total number of words spoken during the meeting. In the bottom panel, we plot this same distribution separately for citizens that live in *Gov* or *Chief* neighborhoods as a fraction of total number of words during meetings from citizens in *Gov* and *Chief* neighborhoods, respectively.

A.3.2 Tax Collection Salience

Table A.11: Effects of Chief Tax Collection: Tax Prime

	Gov meeting		Chief meeting		Either		Both	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Chief tax collection	-0.021 (0.034)	-0.014 (0.033)	-0.025 (0.035)	-0.017 (0.033)	-0.021 (0.035)	-0.014 (0.034)	-0.025 (0.034)	-0.018 (0.033)
Tax prime	-0.013 (0.017)	-0.013 (0.017)	-0.024 (0.017)	-0.024 (0.017)	-0.020 (0.018)	-0.020 (0.017)	-0.018 (0.016)	-0.018 (0.016)
Chief X Tax prime	0.015 (0.027)	0.014 (0.027)	0.029 (0.027)	0.028 (0.027)	0.027 (0.029)	0.026 (0.028)	0.017 (0.026)	0.016 (0.026)
Control mean	0.171	0.171	0.169	0.169	0.180	0.180	0.161	0.161
Observations	2631	2629	2631	2629	2631	2629	2631	2629
Clusters	101	101	101	101	101	101	101	101
Stratum FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Set I	No	Set I	No	Set I	No	Set I

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Coefficients on *Chief tax collection* are OLS estimates with robust standard errors clustered at the polygon level, the unit of randomization and of analysis. *Chief Meeting* is the proportion of citizens requesting a meeting with the chief in a given polygon. *Gov Meeting* is the proportion of citizens requesting a meeting with the provincial government in a given polygon. *Either* is the proportion of citizens in a given polygon requesting either a meeting with the chief or the provincial government. *Both* is the proportion of citizens in a given polygon requesting both a meeting with the chief and with the provincial government. *Chief tax collection* is a binary indicator for whether the polygon experienced tax collection by the chief (1) or by the provincial government (0; control). *Tax prime* is a cross-randomized treatment that indicates whether the respondent had been addressed as a taxpayer (1) on the submission forms or as a citizen (0). Here we compare polygons that had a collective action opportunity ($j = 101$) and where either tax collection was done exclusively by the chief ($j = 51$) or by the provincial government ($j = 50$). *Stratum FE* refer to the stratum used for randomization which is a combination of geographic location of the neighborhood as well as previous tax compliance. *Set I* controls include house type, wave of collective action campaign, distances to letter boxes and polygon-level averages of participation and per capita tax revenue 2016.

Table A.12: Absolute and Relative (Row) Frequencies of Form Submissions by Treatment Groups

	<i>Chief meeting</i>			<i>Government meeting</i>		
<i>Treatment</i>	No	Yes	Row total	No	Yes	Row total
State	1089 (82.9%)	225 (17.1%)	1314	1092 (83.1%)	222 (16.9%)	1314
Chief	1096 (83.2%)	221 (16.8%)	1317	1104 (83.8%)	213 (16.2%)	1317
Column total	2185 (83.0%)	446 (17.0%)	2631	2196 (83.5%)	435 (16.5%)	2631

Table A.13: Effects of Chief Tax Collection: Amount of Taxes Collected

	Gov meeting		Chief meeting		Either		Both	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Chief tax collection	0.014 (0.040)	0.003 (0.043)	0.026 (0.042)	0.017 (0.044)	0.025 (0.042)	0.013 (0.044)	0.015 (0.040)	0.007 (0.043)
Tax revenue	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Chief X Tax revenue	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Control mean	0.169	0.169	0.171	0.171	0.180	0.180	0.161	0.161
Observations	2631	2629	2631	2629	2631	2629	2631	2629
Clusters	101	101	101	101	101	101	101	101
Stratum FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Set I	No	Set I	No	Set I	No	Set I

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Coefficients on *Chief tax collection* are OLS estimates with robust standard errors clustered at the polygon level, the unit of randomization and of analysis. *Chief Meeting* is the proportion of citizens requesting a meeting with the chief in a given polygon. *Gov Meeting* is the proportion of citizens requesting a meeting with the provincial government in a given polygon. *Either* is the proportion of citizens in a given polygon requesting either a meeting with the chief or the provincial government. *Both* is the proportion of citizens in a given polygon requesting both a meeting with the chief and with the provincial government. *Chief tax collection* is a binary indicator for whether the polygon experienced tax collection by the chief (1) or by the provincial government (0; control). Here we compare polygons that had a collective action opportunity ($j = 101$) and where either tax collection was done exclusively by the chief ($j = 51$) or by the provincial government ($j = 50$). *Stratum FE* refer to the stratum used for randomization which is a combination of geographic location of the neighborhood as well as previous tax compliance. *Set I* controls include house type, wave of collective action campaign, distances to letter boxes and polygon-level averages of participation and per capita tax revenue in the 2016 campaign.

A.3.3 Chief Tax Collection Targeting

Table A.14: Effects of Chief Tax Collection on Ethnic Salience

	Closeness to own tribe		Closeness to another tribe	
	(1)	(2)	(3)	(4)
Chief tax collection	0.010 (0.115)	−0.004 (0.119)	−0.028 (0.093)	−0.032 (0.096)
Control mean	3.617	3.617	3.563	3.563
Observations	949	948	952	951
Clusters	101	101	101	101
Stratum FE	Yes	Yes	Yes	Yes
Controls	No	Set I	No	Set I
Sample	C + L	C + L	C + L	C + L

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Coefficients on *Chief tax collection* are OLS estimates with robust standard errors clustered at the polygon level, the unit of randomization and of analysis. *Closeness to own tribe* ranges 0-5 and measures how close a person feels to people on their avenue from the same tribe. *Closeness to different tribe* ranges 0-5 and measures how close a person feels to people on their avenue from another tribe. Outcome variables come from Endline I. *Chief tax collection* is a binary indicator for whether the polygon experienced tax collection by the chief (1) or by the provincial government (0; control). Here we compare polygons that had a collective action opportunity ($j = 101$) and where either tax collection was done exclusively by the chief ($j = 51$) or by the provincial government ($j = 50$). *Stratum FE* refer to the stratum used for randomization which is a combination of geographic location of the neighborhood as well as previous tax compliance. *Set I* controls include house type, wave of collective action campaign, distances to letter boxes and polygon-level averages of participation and per capita tax revenue in 2016.

Table A.15: Effects of Chief Tax Collection: Main Tribe

	Gov meeting		Chief meeting		Either		Both	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Chief tax collection	-0.022 (0.073)	0.003 (0.074)	-0.046 (0.072)	-0.019 (0.073)	-0.039 (0.072)	-0.011 (0.073)	-0.029 (0.073)	-0.005 (0.073)
Main tribe	-0.019 (0.043)	-0.017 (0.045)	-0.023 (0.043)	-0.021 (0.047)	-0.022 (0.041)	-0.021 (0.045)	-0.019 (0.043)	-0.018 (0.046)
Chief X Main tribe	-0.000 (0.065)	-0.007 (0.064)	0.040 (0.065)	0.033 (0.065)	0.025 (0.065)	0.017 (0.064)	0.014 (0.065)	0.009 (0.064)
Control mean	0.169	0.169	0.171	0.171	0.180	0.180	0.161	0.161
Observations	1339	1339	1339	1339	1339	1339	1339	1339
Clusters	98	98	98	98	98	98	98	98
Stratum FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Set I	No	Set I	No	Set I	No	Set I

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Coefficients on *Chief tax collection* are OLS estimates with robust standard errors clustered at the polygon level, the unit of randomization and of analysis. *Chief Meeting* is the proportion of citizens requesting a meeting with the chief in a given polygon. *Gov Meeting* is the proportion of citizens requesting a meeting with the provincial government in a given polygon. *Either* is the proportion of citizens in a given polygon requesting either a meeting with the chief or the provincial government. *Both* is the proportion of citizens in a given polygon requesting both a meeting with the chief and with the provincial government. *Chief tax collection* is a binary indicator for whether the polygon experienced tax collection by the chief (1) or by the provincial government (0; control). *Main tribe* indicates whether the respondent identifies with the main tribe or not. Here we compare polygons that had a collective action opportunity ($j = 101$) and where either tax collection was done exclusively by the chief ($j = 51$) or by the provincial government ($j = 50$). *Stratum FE* refer to the stratum used for randomization which is a combination of geographic location of the neighborhood as well as previous tax compliance. *Set I* controls include house type, wave of collective action campaign, distances to letter boxes and polygon-level averages of participation and per capita tax revenue in 2016.

A.4 Exploratory Heterogeneous Effects

Table A.16: Effects of Chief Tax Collection: Respondent Gender

	Gov meeting		Chief meeting		Either		Both	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Chief tax collection	-0.021 (0.051)	-0.006 (0.048)	-0.014 (0.052)	0.001 (0.049)	-0.019 (0.053)	-0.004 (0.050)	-0.016 (0.051)	-0.001 (0.048)
Woman	0.030 (0.034)	0.026 (0.031)	0.012 (0.038)	0.009 (0.035)	0.025 (0.037)	0.022 (0.035)	0.018 (0.035)	0.013 (0.032)
Chief X Woman	0.006 (0.055)	0.015 (0.052)	0.009 (0.060)	0.018 (0.057)	0.011 (0.060)	0.020 (0.057)	0.004 (0.055)	0.013 (0.052)
Control mean	0.169	0.169	0.171	0.171	0.180	0.180	0.161	0.161
Observations	1570	1570	1570	1570	1570	1570	1570	1570
Clusters	98	98	98	98	98	98	98	98
Stratum FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Set I	No	Set I	No	Set I	No	Set I

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Coefficients on *Chief tax collection* are OLS estimates with robust standard errors clustered at the polygon level, the unit of randomization and of analysis. *Chief Meeting* is the proportion of citizens requesting a meeting with the chief in a given polygon. *Gov Meeting* is the proportion of citizens requesting a meeting with the provincial government in a given polygon. *Either* is the proportion of citizens in a given polygon requesting either a meeting with the chief or the provincial government. *Both* is the proportion of citizens in a given polygon requesting both a meeting with the chief and with the provincial government. *Chief tax collection* is a binary indicator for whether the polygon experienced tax collection by the chief (1) or by the provincial government (0; control). Here we compare polygons that had a collective action opportunity ($j = 101$) and where either tax collection was done exclusively by the chief ($j = 51$) or by the provincial government ($j = 50$). *Stratum FE* refer to the stratum used for randomization which is a combination of geographic location of the neighborhood as well as previous tax compliance. *Set I* controls include house type, wave of collective action campaign, distances to letter boxes and polygon-level averages of participation and per capita tax revenue in the 2016 campaign.

A.5 Discussion of Differences with Pre-Analysis Plan

We show the results of further analyses included in the pre-analysis plan. We first document that chief tax collection did not lead to statistically detectable increases in trust in institutions (Table A.17), although the coefficients are non-trivial. We then analyze additional endline I and endline II outcomes in Tables A.18 and A.19. Finally, we compare participation in chief and state tax collection neighborhoods to five neighborhoods in which collectors only went to deliver tax bills but never returned to collect taxes; citizens were expected to pay themselves at the tax ministry. With only five neighborhoods, this analysis is unfortunately under-powered but for completeness we present the results in Table A.20.

Table A.17: Effects of Chief Tax Collection on Trust in Institutions

	Trust in			
	Gov	Gov	Chief	Chief
Chief tax collection	0.123 (0.102)	0.144 (0.099)	0.104 (0.088)	0.103 (0.081)
Control mean	-0.068	-0.068	-0.067	-0.067
Observations	958	957	946	945
Clusters	101	101	101	101
Stratum FE	Yes	Yes	Yes	Yes
Controls	No	Set I	No	Set I

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Coefficients on *Chief tax collection* are OLS estimates with robust standard errors clustered at the polygon level, the unit of randomization and of analysis. Outcomes in columns 1-4 come from the Endline I survey and have been z-transformed such that coefficients represent standardized mean differences. *Trust* measures citizens confidence in the work of the chief and the government, respectively. *Chief tax collection* is a binary indicator for whether the polygon experienced tax collection by the chief (1) or by the provincial government (0; control). Here we compare polygons that had a collective action opportunity ($j = 101$) and where either tax collection was done exclusively by the chief ($j = 51$) or by the provincial government ($j = 50$). We restrict the sample here further to exclude outliers on the capacity questions as well as villas. *Stratum FE* refer to the stratum used for randomization which is a combination of geographic location of the neighborhood as well as previous tax compliance. *Set I* controls include house type, wave of collective action campaign, distances to letter boxes and polygon-level averages of participation and per capita tax revenue in 2016.

Table A.18: Effects of Chief Tax Collection on Additional Outcomes in Endline I

	Evaluation of		Corruption by		Capacity of	
	Chief	Chief	Chief	Chief	Gov	Gov
Chief tax collection	-0.007 (0.089)	-0.003 (0.083)	-26.324 (26.111)	-29.406 (27.048)	-0.018 (0.040)	-0.012 (0.040)
Control mean	3.487	3.487	463.517	463.517	0.554	0.554
Observations	675	674	743	742	963	962
Clusters	96	96	99	99	101	101
Stratum FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Set I	No	Set I	No	Set I

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Coefficients on *Chief tax collection* are OLS estimates with robust standard errors clustered at the polygon level, the unit of randomization and of analysis. Outcomes in columns 1-4 come from the Endline I survey. *Evaluation of chief* measures citizens' perception of the performance of the chief. *Corruption by chief* measures citizens' evaluation of how much of a public works project budget would be stolen by the chief. *Capacity of gov* measures citizens' perception of the capacity of the state to repair roads within three months. *Chief tax collection* is a binary indicator for whether the polygon experienced tax collection by the chief (1) or by the provincial government (0; control). Here we compare polygons that had a collective action opportunity ($j = 101$) and where either tax collection was done exclusively by the chief ($j = 51$) or by the provincial government ($j = 50$). We restrict the sample here further to exclude outliers on the capacity questions as well as villas. *Stratum FE* refer to the stratum used for randomization which is a combination of geographic location of the neighborhood as well as previous tax compliance. *Set I* controls include house type, wave of collective action campaign, distances to letter boxes and polygon-level averages of participation and per capita tax revenue in 2016.

Table A.19: Effects of Chief Tax Collection on Additional Outcomes in Endline II

	Gov eval		DIVAS eval		Chief responsive	
	(1)	(2)	(3)	(4)	(5)	(6)
Chief tax collection	-0.073 (0.121)	-0.022 (0.103)	-0.111 (0.070)	-0.100 (0.070)	0.057 (0.067)	0.088 (0.064)
Control mean	4.124	4.124	4.583	4.583	2.982	2.982
Observations	1191	1191	1077	1077	1427	1427
Clusters	101	101	100	100	99	99
Stratum FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Set I	No	Set I	No	Set I
Sample	C + L	C + L	C + L	C + L	C + L	C + L

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Coefficients on *Chief tax collection* are OLS estimates with robust standard errors clustered at the polygon level, the unit of randomization and of analysis. *Chief Meeting* is the proportion of citizens requesting a meeting with the chief in a given polygon. Outcomes come from the Endline II survey. *Gov eval* measures citizens' evaluations of the performance of the provincial government in Kananga. *DIVAS eval* measures citizens' evaluations of the performance of the Division des Affaires Sociales (DIVAS). *Chief responsive* measures citizens' perception of the degree to which the chief responds to the needs of the jurisdiction's inhabitants. *Chief tax collection* is a binary indicator for whether the polygon experienced tax collection by the chief (1) or by the provincial government (0; control). Here we compare polygons that had a collective action opportunity ($j = 101$) and where either tax collection was done exclusively by the chief ($j = 51$) or by the provincial government ($j = 50$). We restrict the sample here further to exclude compounds classified as villas. *Stratum FE* refer to the stratum used for randomization which is a combination of geographic location of the neighborhood as well as previous tax compliance. *Set I* controls include house type, wave of collective action campaign, distances to letter boxes and polygon-level averages of participation and per capita tax revenue in 2016.

Table A.20: Effects of Chief and Govt. Collection on Demand for Citizen Meetings: Comparison to Pure Control Group

	Gov meeting				Chief meeting			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Chief tax collection	-0.081 (0.095)	-0.057 (0.097)			-0.051 (0.080)	-0.030 (0.084)		
Gov tax collection			-0.074 (0.096)	-0.075 (0.095)			-0.048 (0.081)	-0.044 (0.084)
Control mean	0.243	0.243	0.243	0.243	0.219	0.219	0.219	0.219
Observations	1486	1486	1483	1481	1486	1486	1483	1481
Clusters	56	56	55	55	56	56	55	55
FE	None	None	None	None	None	None	None	None
Controls	None	Set I	None	Set I	None	Set I	None	Set I

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Coefficients on *Chief tax collection* are OLS estimates with robust standard errors clustered at the polygon level, the unit of randomization and of analysis. *Chief Meeting* is the proportion of citizens requesting a meeting with the chief in a given polygon. *Gov Meeting* is the proportion of citizens requesting a meeting with the provincial government in a given polygon. *Either* is the proportion of citizens in a given polygon requesting either a meeting with the chief or the provincial government. *Both* is the proportion of citizens in a given polygon requesting both a meeting with the chief and with the provincial government. *Chief tax collection* is a binary indicator for whether the polygon experienced tax collection by the chief (1) or whether the polygon was in the Pure Control group (0). *Gov tax collection* is a binary indicator for whether the polygon experienced tax collection by the government (1) or whether the polygon was in the Pure Control group (0). Here we compare polygons that had a collective action opportunity ($J = 106$) with treatment arms govt. collection, chief collection or the pure control group. In this analysis, we do not include fixed effects for randomization strata since we do not achieve sufficient saturation of each stratum with at least one pure control polygon. *Set I* controls include house type, wave of collective action campaign, distances to letter boxes and polygon-level averages of participation and per capita tax revenue in the 2016 campaign.

Appendix B

Do Anti-Corruption Policies Attract Votes? Experimental Evidence from India and the U.K.

B.1 Additional Information on Experimental Design

B.1.1 UK Survey Instrument

Q1: In what year were you born?

Q2: Which best describes your gender?

Male

Female

Other

Q3: What is the highest level of education you have completed?

GCSEs/ O-Levels

A-Levels

Bachelor's degree

Trade/Technical/Vocational Training

Postgraduate degree

Q4: What was your total household income in the previous year?

Less than £10,000

£10,000 - £19,999

£20,000 - £39,999

£40,000 - £59,999

£60,000 - £99,999

£100,000 and over

Q5: When people talk about politics, the terms "left" and "right" are always used. Where would you place yourself on this scale, where 1 means left and 10 means right?

Q6: Is there a political party with which you identify?

Yes

No

[Answer Q6.1 and Q6.2 if "Yes" is selected.]

Q6.1: With which political party do you identify?

Labour

Conservative

Liberal Democrat

Scottish National Party (SNP)

Plaid Cymru

Green Party

United Kingdom Independence Party (UKIP)

Other

None

Q6.2: Do you identify very strongly or not very strongly with this party?

Very strongly

Not very strongly

What follows is a series of choice tasks, where you are presented with 6 tasks in total. Please read the descriptions of the following political candidates carefully. They propose policy packages and have certain candidate characteristics. Then, please indicate which of the two candidates you would personally prefer to see elected in the next election.

Endline survey: Outcome measure

Q7: If you had to choose between them, which of these two candidates would you vote for in the upcoming election?

Table B.1: UK Conjoint: Attribute Levels

Attribute	Level	Text in table
Costs	Public money/taxes	Raising income taxes
	Opportunity costs	Less money for other issues
	Personal data	Disclosure of private data
	No costs	No costs
Effective	Effective	Has been shown by experts to perform well
	Not effective	Has been shown by experts to have little effect
Issue	Anti-corruption	Anti-corruption
	Economy	Economy
	Domestic Security	Domestic Security
Credibility	Credible	Has a reputation for keeping campaign promises
	Not credible	Has a reputation for not keeping campaign promises
Gender	Female	Female
	Male	Male
Party	Incumbent party	Conservative
	Main opposition party	Labour

Candidate 1

Candidate 2

Q8: On a scale from 1 to 7, where 1 indicates absolutely no support for the candidate and 7 the highest support, how would you rate candidate 1?

- 1
- 2
- 3
- 4
- 5
- 6
- 7

Q9: Using the same scale, how would you rate candidate 2?

- 1
- 2
- 3
- 4
- 5
- 6
- 7

Q10: In the six choice tasks you've just completed, how many attributes did each candidate have?

- 2
- 4
- 6
- 8

Q11: Imagine you are planning the federal budget for the upcoming budgetary year. Please allocate percentages of money that should be spent on each issue. The percentages must add up to 100%.

- Anti-corruption
- Economic prosperity
- Domestic security
- Environment
- Education
- Social equality

Q12: Have you ever received material benefits in exchange for your voting for a particular candidate?

- Yes
- No
- Do not wish to answer

B.1.2 India Survey Instrument

Q1: In what year were you born?

Q2: Which best describes your gender?

Male

Female

Other

Q3: What is the highest level of education you have completed?

General Secondary School (Standard X)

Upper Secondary School (Standard XII)

Trade/Technical/Vocational Training

Bachelor's degree

Postgraduate degree

Q4: What was your total household income in the previous year?

Less than 60,000

60,000 - 89,999

90,000 - 119,999

120,000 - 239,999

240,000 - 599,999

600,000 and over

Q5: When people talk about politics, the terms "left" and "right" are always used.

Where would you place yourself on this scale, where 1 means left and 10 means right?

Q6: Is there a political party with which you identify?

Yes

No

[Answer Q6.1 and Q6.2 if "Yes" is selected.]

Q6.1: With which political party do you identify?

Bharatiya Janata Party

Indian National Congress

Bahujan Samaj Party

All India Trinamool Congress

Samajwadi Party

Other

Q6.2: Do you identify very strongly or not very strongly with this party?

Very strongly

Not very strongly

What follows now is a series of choice tasks, where you are presented with 6 tasks in total.

Please read the descriptions of the following political candidates carefully. They propose policy packages and have certain candidate characteristics. Then, please indicate which of the two candidates you would personally prefer to see elected in the next election.

Table B.2: India Conjoint: Attribute Levels

Attribute	Level	Text in table
Costs	Public money/taxes	Raising income taxes
	Opportunity costs	Less money for other issues
	Personal data	Disclosure of private data
	No costs	No costs
Effective	Effective	Has been shown by experts to perform well
	Not effective	Has been shown by experts to have little effect
Issue	Anti-corruption	Anti-corruption
	Economy	Economy
	Domestic Security	Domestic Security
Credibility	Credible	Has a reputation for keeping campaign promises
	Not credible	Has a reputation for not keeping campaign promises
Gender	Female	Female
	Male	Male
Party	Incumbent party	Indian National Congress
	Main opposition party	Bharatiya Janata Party

Endline survey: Outcome measure

Q7: If you had to choose between them, which of these two candidates would you vote for in the upcoming election?

Candidate 1

Candidate 2

Q8: On a scale from 1 to 7, where 1 indicates absolutely no support for the candidate and 7 the highest support, how would you rate candidate 1?

- 1
- 2
- 3
- 4
- 5
- 6
- 7

Q9: Using the same scale, how would you rate candidate 2?

- 1
- 2
- 3
- 4
- 5
- 6
- 7

Q10: In the six choice tasks you've just completed, how many attributes did each candidate have?

- 2
- 4
- 6
- 8

Q11: Imagine you are planning the federal budget for the upcoming budgetary year. Please allocate percentages of money that should be spent on each issue. The percentages must add up to 100%.

Anti-corruption

Economic prosperity

Domestic security

Environment

Education

Social equality

Q12: Have you ever received material benefits in exchange for your voting for a particular candidate?

Yes

No

Do not wish to answer

B.1.3 UK Subject Sample Characteristics

Table B.3: UK Subject Sample Characteristics

Variable	$\bar{x}$	X_0	X_{100}	N
Age	43.9	18	80	193
Female	0.51	0	1	196
L-R self	4.6	1	10	196
GCSEs	0.16	0	1	192
A-Levels	0.20	0	1	192
Voc. Training	0.12	0	1	192
Bachelor's	0.33	0	1	192
Postgrad	0.20	0	1	192
< £10,000	0.12	0	1	192
£10,000 - £19,999	0.17	0	1	192
£20,000 - £39,999	0.37	0	1	192
£40,000 - £59,999	0.17	0	1	192
£60,000 - £99,999	0.10	0	1	192
> £100,000	0.06	0	1	192

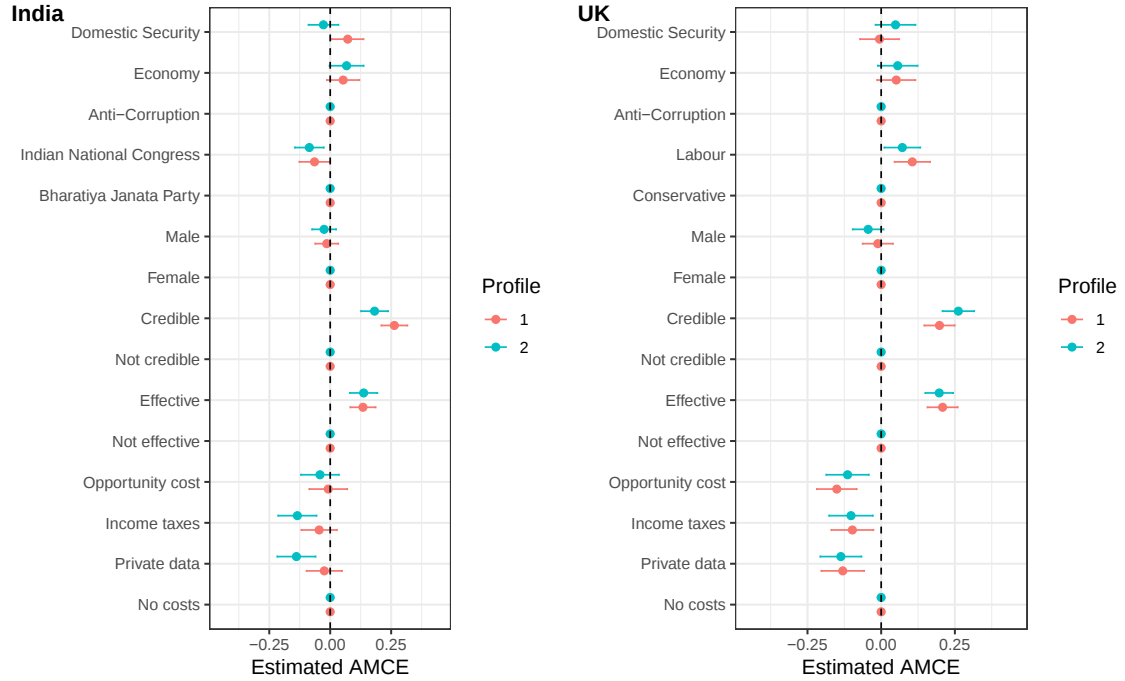
B.1.4 India Subject Sample Characteristics

Table B.4: India Subject Sample Characteristics

Variable	$\bar{x}$	X_0	X_{100}	N
Age	28.4	17	65	205
Female	0.28	0	1	212
L-R self	6.2	1	10	212
Gen. Sec. School	0.01	0	1	207
Upper Sec. School	0.18	0	1	207
Voc. Training	0.02	0	1	207
Bachelor's	0.42	0	1	207
Postgrad	0.38	0	1	207
< 60,000	0.10	0	1	165
60,000 - 89,999	0.07	0	1	165
90,000 - 119,999	0.09	0	1	165
120,000 - 239,999	0.08	0	1	165
240,000 - 599,999	0.27	0	1	165
$\geq$ 600,000	0.37	0	1	165

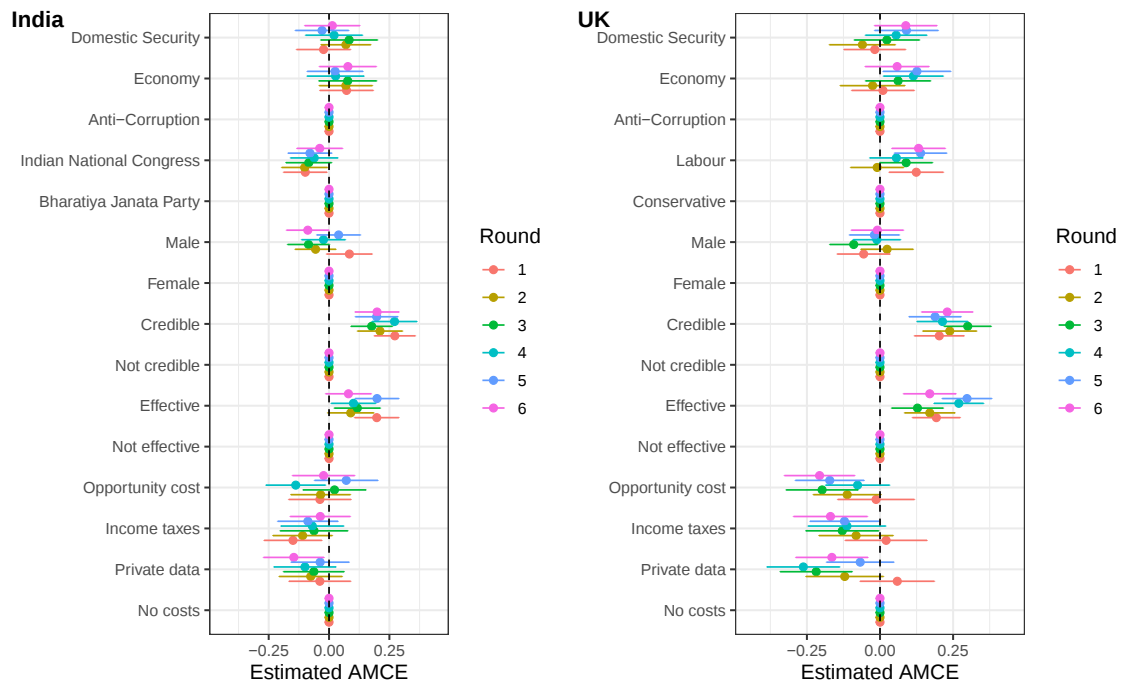
B.1.5 Design Diagnostics: Profile Order and Carryover Effects

Figure B.1: Test for Profile Order Effects



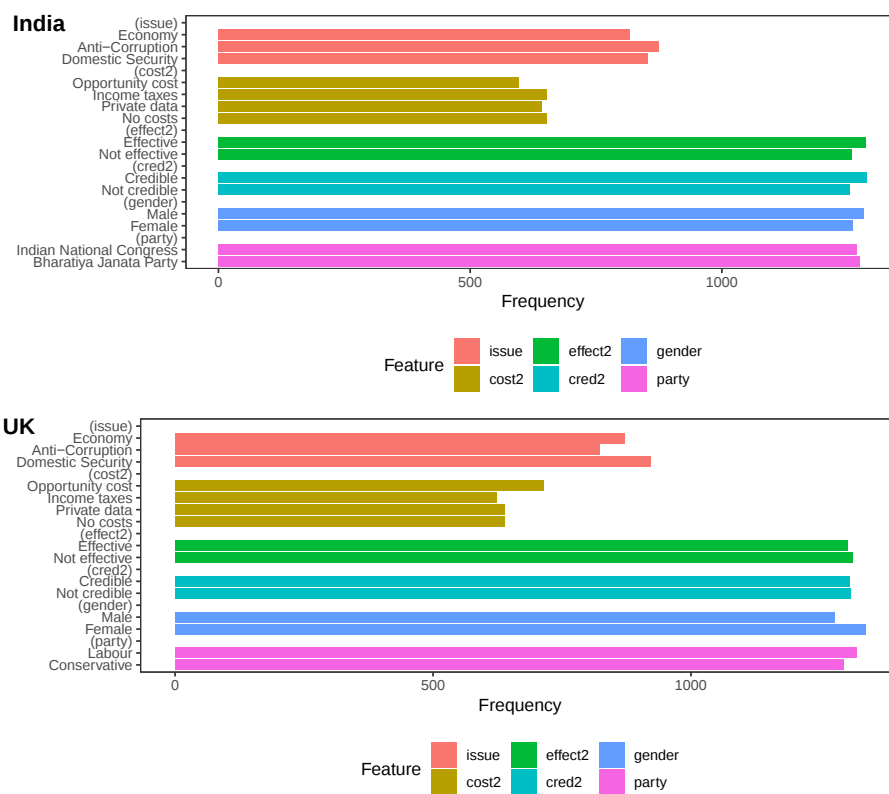
Notes: This figure displays estimated Average Marginal Component Effects (AMCEs) for the candidate choice experiment in India (left panel) and the U.K. (right panel) conditional on profile order. These are results from binary logit models where the sample is subset to strata defined by whether the attribute level appeared in the first or second profile and the outcome, candidate choice, is regressed on series of categorical variables indicating policy issues, costs and effectiveness as well as candidate credibility, gender, and party. 95% confidence intervals are based on standard errors clustered at the respondent level. All analyses are implemented with the R package `cregg`.

Figure B.2: Test for Carryover Effects



Notes: This figure displays estimated Average Marginal Component Effects (AMCEs) for the candidate choice experiment in India (left panel) and the U.K. (right panel) conditional on conjoint round. These are results from binary logit models where the sample is subset to strata defined by which round the attribute level appeared in and the outcome, candidate choice, is regressed on series of categorical variables indicating policy issues, costs and effectiveness as well as candidate credibility, gender, and party. 95% confidence intervals are based on standard errors clustered at the respondent level. All analyses are implemented with the R package `cregg`.

Figure B.3: Display Frequencies of Conjoint Levels



Notes: This figure displays the absolute frequencies with which each of the attribute levels appeared in the conjoint tasks.

B.2 Appendices for Main Results

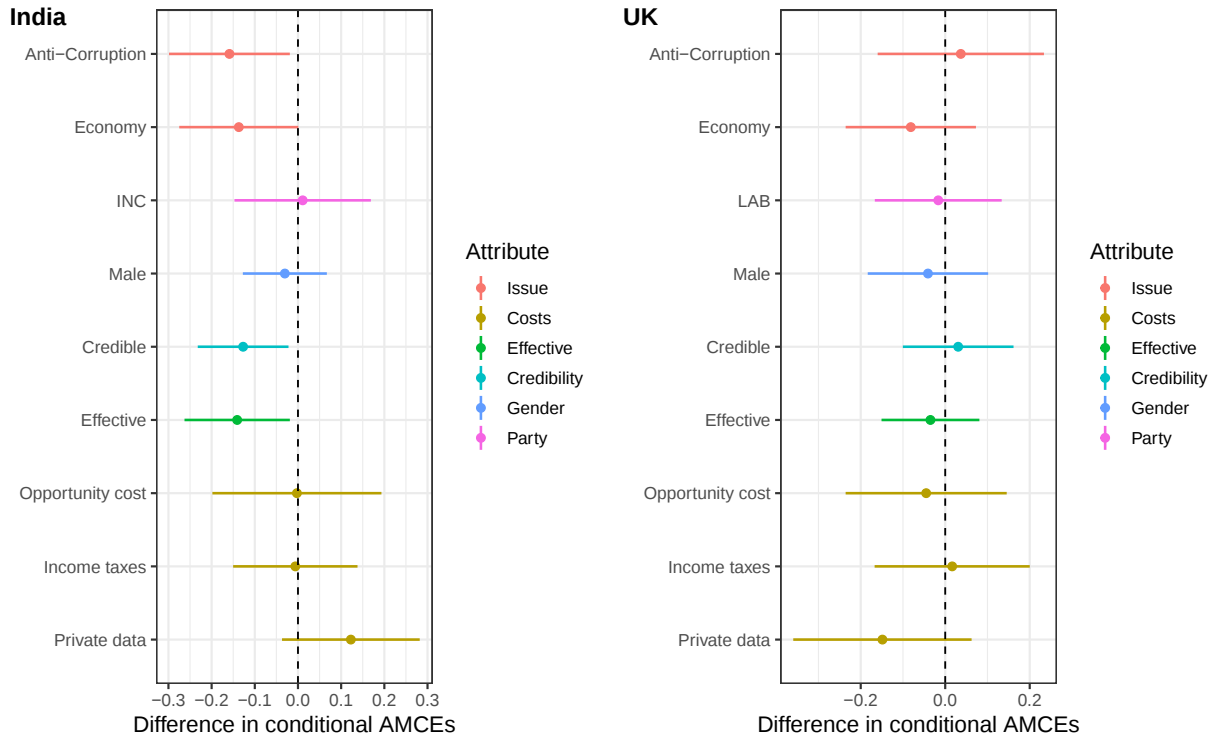
Table B.5: India AMCEs

feature	level	estimate	std.error	p	lower	upper
Issue	Economy	0				
Issue	Domestic Security	-0.037	0.026	0.149	-0.087	0.013
Issue	Anti-Corruption	-0.061	0.026	0.020	-0.112	-0.010
Costs	No costs	0				
Costs	Private data	-0.084	0.027	0.002	-0.136	-0.031
Costs	Income taxes	-0.089	0.027	0.001	-0.142	-0.035
Costs	Opportunity cost	-0.029	0.027	0.290	-0.082	0.025
Effective	Not effective	0				
Effective	Effective	0.132	0.020	0	0.093	0.172
Credibility	Not credible	0				
Credibility	Credible	0.222	0.021	0	0.181	0.262
Gender	Female	0				
Gender	Male	-0.021	0.017	0.222	-0.055	0.013
Party	BJP	0				
Party	INC	-0.076	0.025	0.002	-0.125	-0.027

Table B.6: UK AMCEs

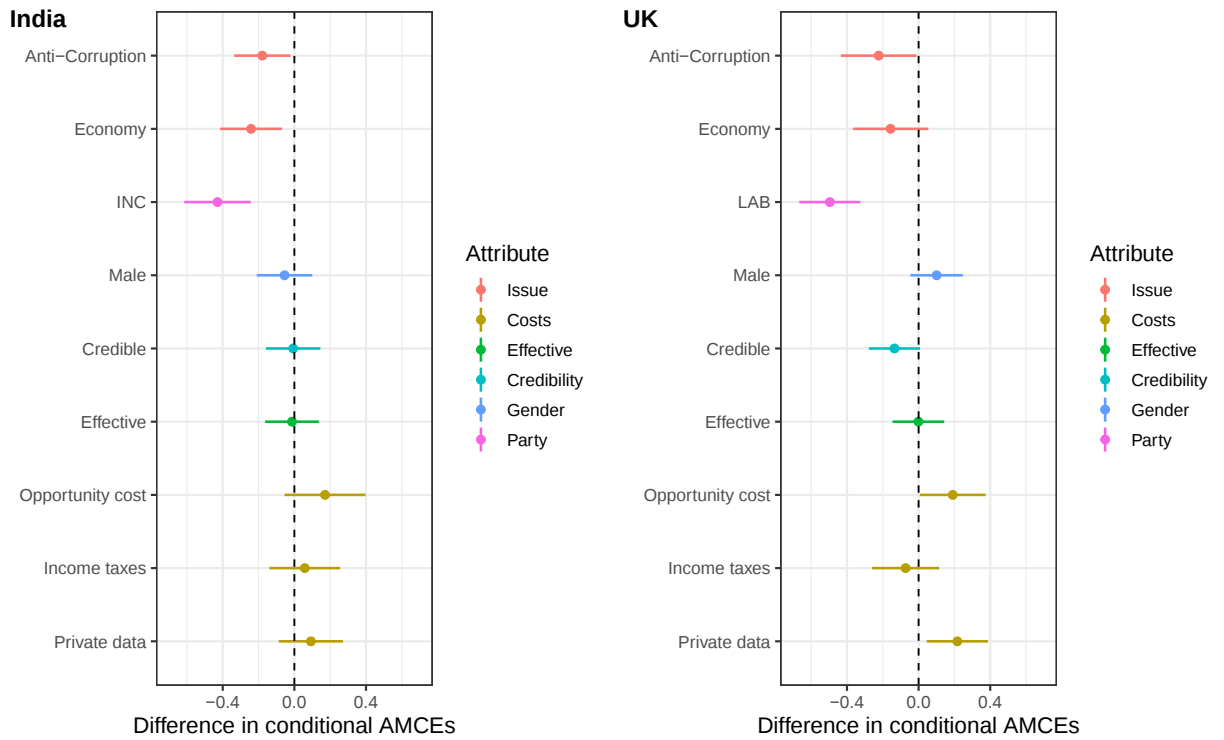
feature	level	estimate	std.error	p	lower	upper
Issue	Economy	0				
Issue	Domestic Security	-0.030	0.024	0.204	-0.077	0.016
Issue	Anti-Corruption	-0.053	0.023	0.019	-0.098	-0.009
Costs	No costs	0				
Costs	Private data	-0.134	0.027	0.00000	-0.188	-0.081
Costs	Income taxes	-0.100	0.027	0.0002	-0.152	-0.048
Costs	Opportunity cost	-0.132	0.026	0.00000	-0.184	-0.080
Effective	Not effective	0				
Effective	Effective	0.204	0.018	0	0.169	0.239
Credibility	Not credible	0				
Credibility	Credible	0.231	0.019	0	0.193	0.269
Gender	Female	0				
Gender	Male	-0.027	0.019	0.151	-0.063	0.010
Party	CON	0				
Party	LAB	0.089	0.024	0.0002	0.042	0.136

Figure B.4: Difference in AMCEs by Income



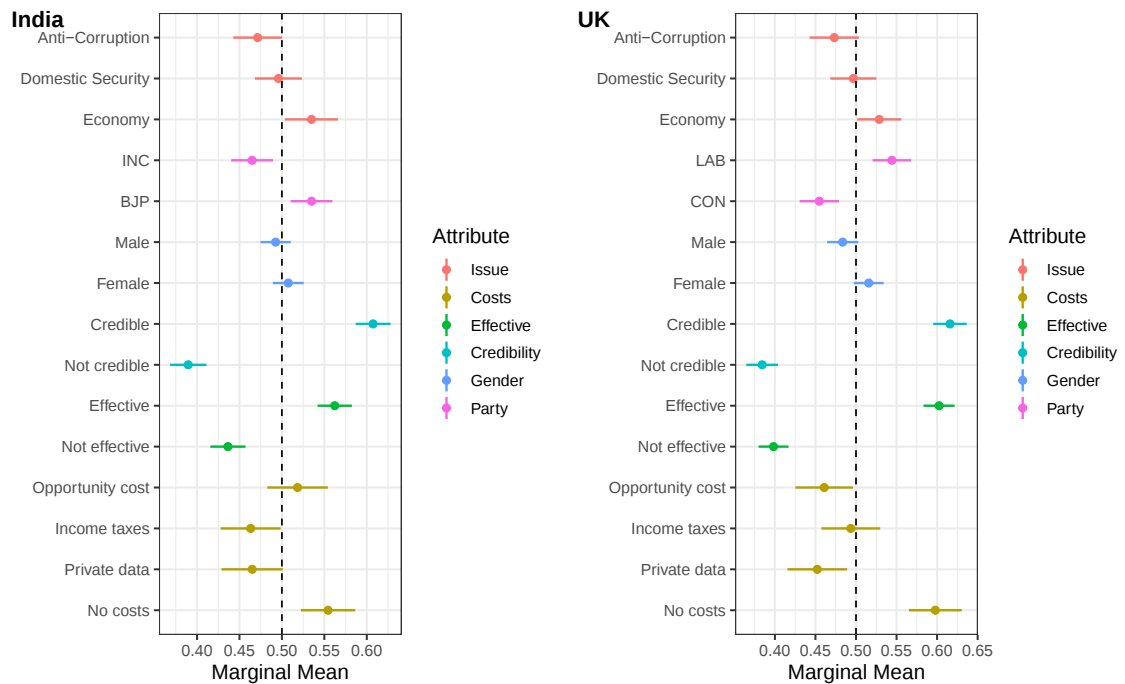
Notes: This figure displays estimated differences in Average Marginal Component Effects (AMCEs) for the candidate choice experiment in India (left panel) and the U.K. (right panel) between low and high levels of household income. These are results from binary logit models where the sample is subset to strata defined by household income and the outcome, candidate choice, is regressed on series of categorical variables indicating policy issues, costs and effectiveness as well as candidate credibility, gender, and party. 95% confidence intervals are based on standard errors clustered at the respondent level. All analyses are implemented with the R package `cregg`.

Figure B.5: Difference in AMCEs by Incumbent Support



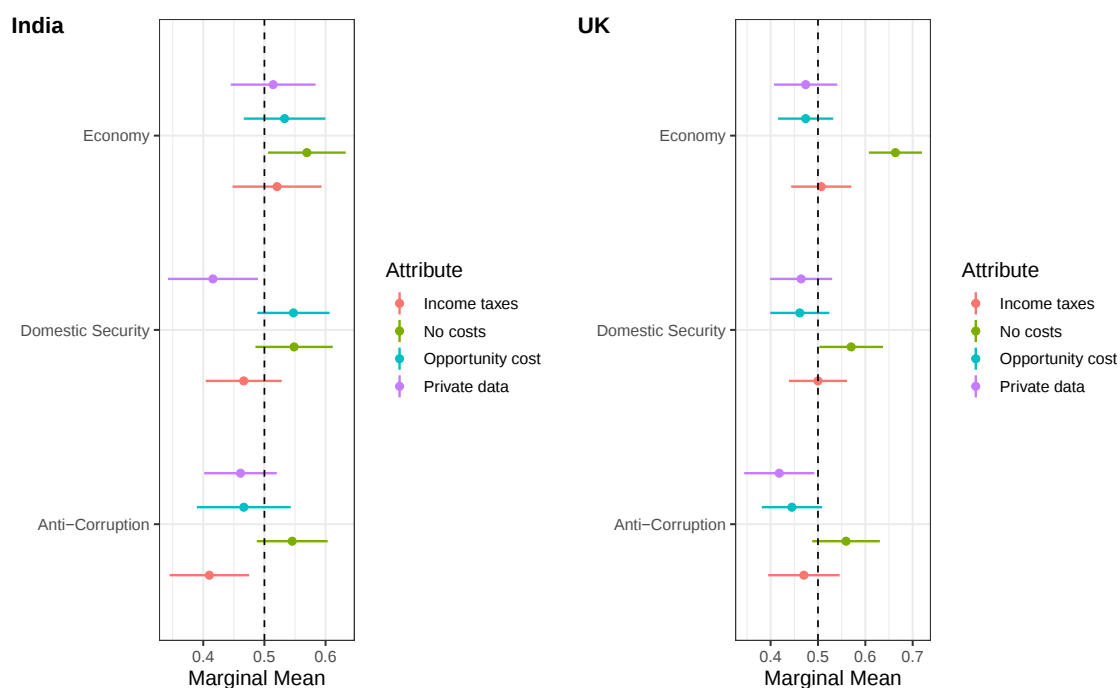
Notes: This figure displays estimated differences in Average Marginal Component Effects (AMCEs) for the candidate choice experiment in India (left panel) and the U.K. (right panel) between incumbent party supporters and all other respondents. These are results from binary logit models where the sample is subset to strata defined by incumbent party support and the outcome, candidate choice, is regressed on series of categorical variables indicating policy issues, costs and effectiveness as well as candidate credibility, gender, and party. 95% confidence intervals are based on standard errors clustered at the respondent level. All analyses are implemented with the R package `cregg`.

Figure B.6: Marginal Means



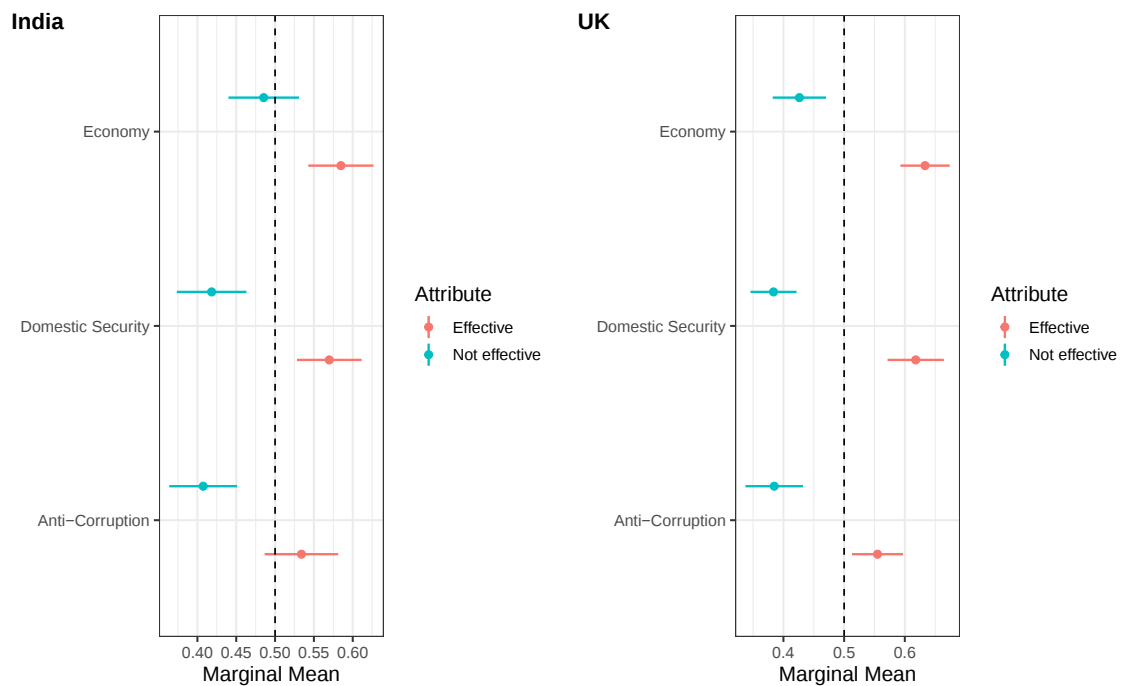
Notes: This figure displays estimated Marginal Means (MMs) for the candidate choice experiment in India (left panel) and the U.K. (right panel). The marginal mean of a conjoint attribute level is the mean outcome across all appearances of a particular conjoint attribute level averaged over all other attributes. In forced-choice designs with two profiles such as the one employed here, MMs by design are on average 0.5 where values above 0.5 indicate increases in profile favorability and values below 0.5 indicate decreasing profile favorability. All analyses are implemented with the R package `cregg`.

Figure B.7: Marginal Means for Issue/Costs Interaction



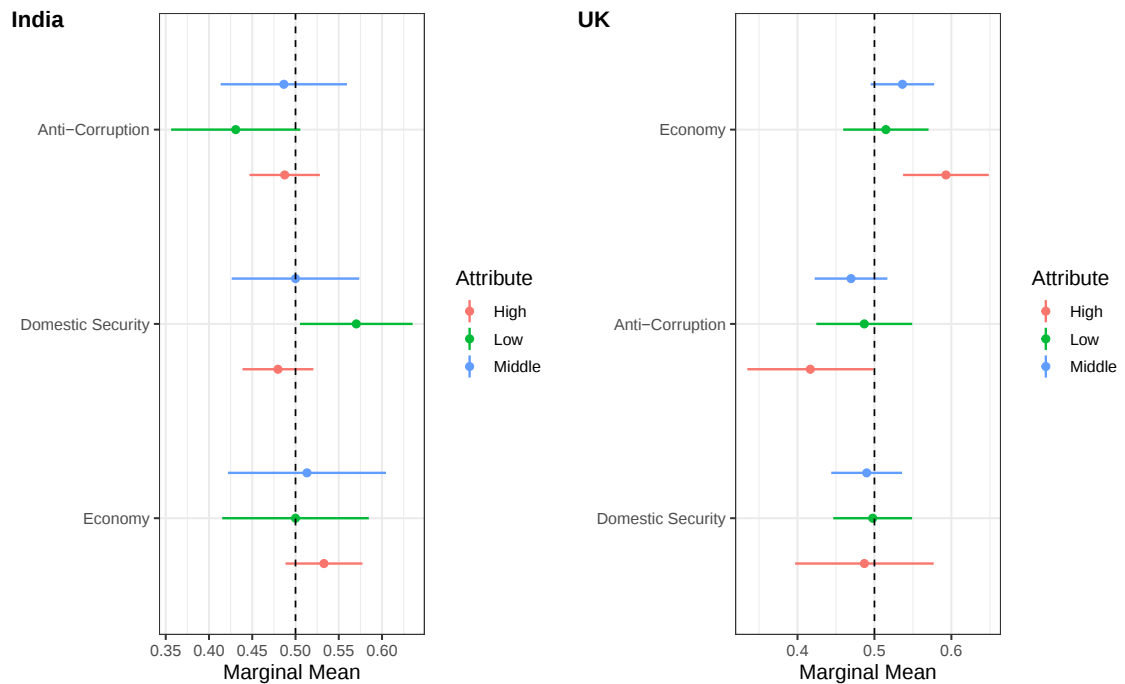
Notes: This figure displays estimated Marginal Means (MMs) for the candidate choice experiment in India (left panel) and the U.K. (right panel) for the interaction between policy issue and costs. The marginal mean of a conjoint attribute level is the mean outcome across all appearances of a particular conjoint attribute level averaged over all other attributes. In forced-choice designs with two profiles such as the one employed here, MMs by design are on average 0.5 where values above 0.5 indicate increases in profile favorability and values below 0.5 indicate decreasing profile favorability. All analyses are implemented with the R package `cregg`.

Figure B.8: Marginal Means for Issue/Effectiveness Interaction



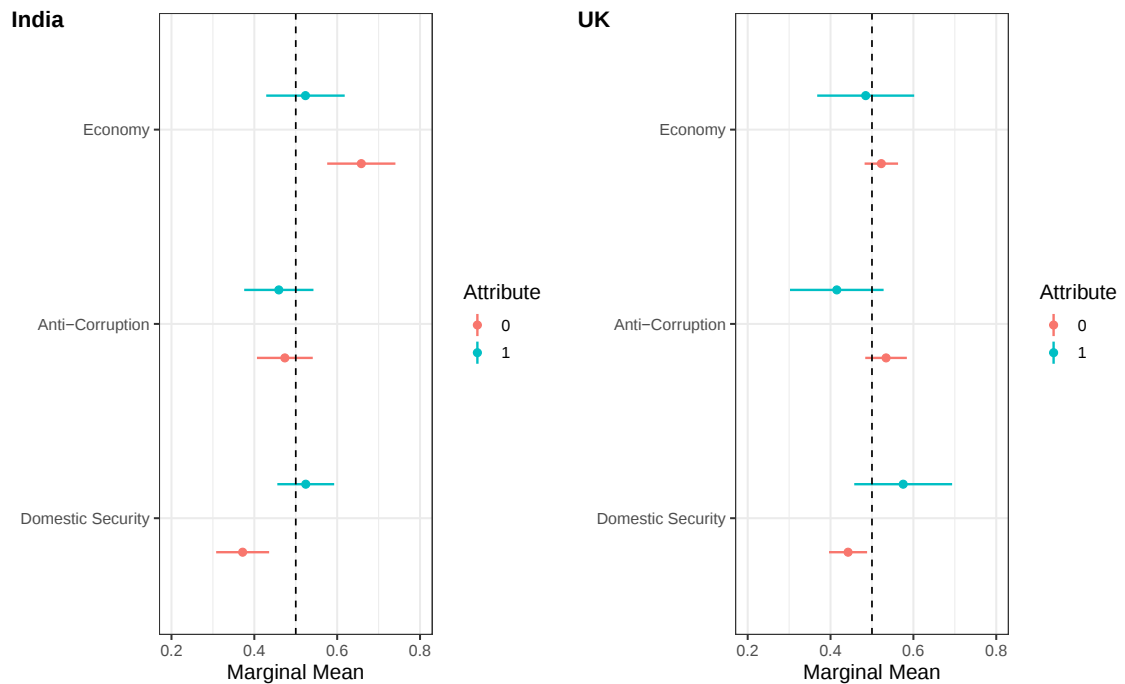
Notes: This figure displays estimated Marginal Means (MMs) for the candidate choice experiment in India (left panel) and the U.K. (right panel) for the interaction between policy issue and effectiveness. The marginal mean of a conjoint attribute level is the mean outcome across all appearances of a particular conjoint attribute level averaged over all other attributes. In forced-choice designs with two profiles such as the one employed here, MMs by design are on average 0.5 where values above 0.5 indicate increases in profile favorability and values below 0.5 indicate decreasing profile favorability. All analyses are implemented with the R package `cregg`.

Figure B.9: Subgroup Marginal Means by Income



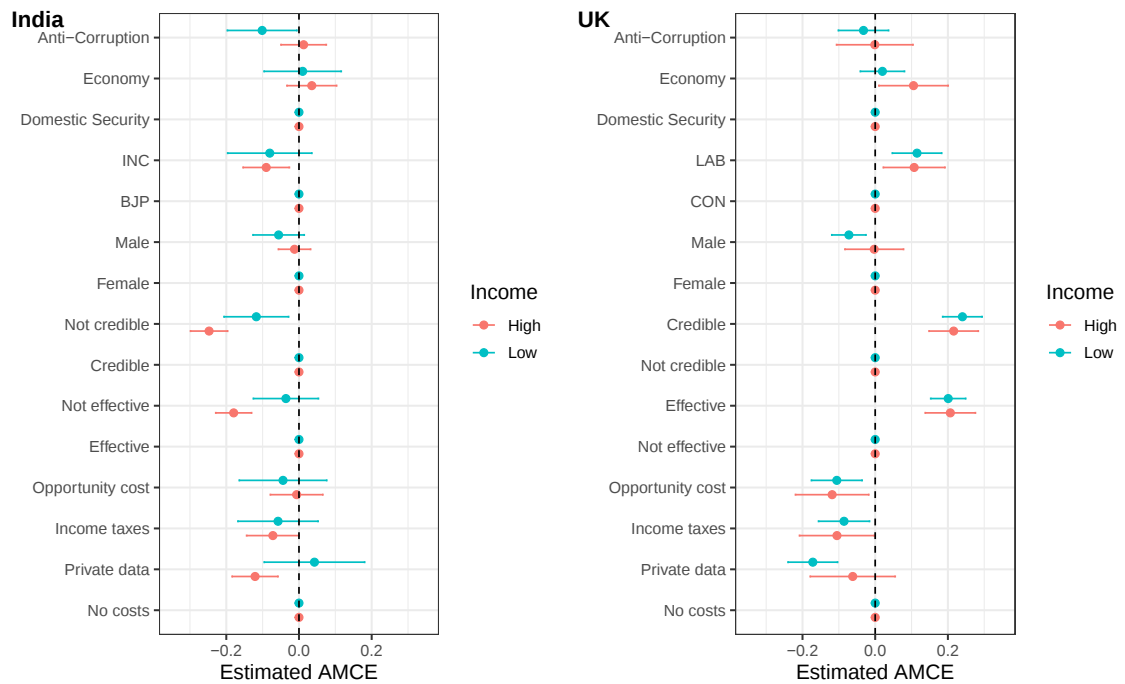
Notes: This figure displays estimated Marginal Means (MMs) for the candidate choice experiment in India (left panel) and the U.K. (right panel) for the interaction between policy issue and household income. The marginal mean of a conjoint attribute level is the mean outcome across all appearances of a particular conjoint attribute level averaged over all other attributes. In forced-choice designs with two profiles such as the one employed here, MMs by design are on average 0.5 where values above 0.5 indicate increases in profile favorability and values below 0.5 indicate decreasing profile favorability. All analyses are implemented with the R package `cregg`.

Figure B.10: Subgroup Marginal Means by Incumbent Support



Notes: This figure displays estimated Marginal Means (MMs) for the candidate choice experiment in India (left panel) and the U.K. (right panel) for the interaction between policy issue and incumbent party support. The marginal mean of a conjoint attribute level is the mean outcome across all appearances of a particular conjoint attribute level averaged over all other attributes. In forced-choice designs with two profiles such as the one employed here, MMs by design are on average 0.5 where values above 0.5 indicate increases in profile favorability and values below 0.5 indicate decreasing profile favorability. All analyses are implemented with the R package `cregg`.

Figure B.11: AMCEs by Household Income: Binary



Notes: This figure displays estimated Average Marginal Component Effects (AMCEs) for the candidate choice experiment in India (left panel) and the U.K. (right panel) conditional on household income. These are results from binary logit models where the sample is subset to strata defined by a binary discretization of household income and the outcome, candidate choice, is regressed on series of categorical variables indicating policy issues, costs and effectiveness as well as candidate credibility, gender, and party. 95% confidence intervals are based on standard errors clustered at the respondent level. All analyses are implemented with the R package `cregg`.

B.3 Discussion of Differences with Pre-Analysis Plan

This study was pre-registered informally on the author’s github page in December 2018 from which the PAP version in this Appendix is drawn and the upload date and version of the PAP can be verified via version control on github.

It is important to note a few deviations from this original pre-analysis plan. I list these deviations below.

- Framing and theory outlined in the original pre-analysis plan have changed substantially and therefore hypotheses have been slightly modified to reflect those changes
- From the two experiments outlined in the pre-analysis plan (1 standard vignette and 1 conjoint), only the conjoint experiment was finally run due to funding constraints
- A third hypothesized conditional AMCE (individual experience with clientelism/corruption) was not included in the final analysis due to concerns about nonresponse and misreporting bias

Appendix C

Tax Earmarking and Political Participation: Theory and Evidence from Ghana

C.1 Additional Information on Experimental Design

C.1.1 Survey Instrument

Q1: How old are you? If you prefer not to say, please write “Prefer not to say”.

Q2: What is your gender?

Male

Female

Prefer not to say

Q3: What is your highest level of education?

No formal schooling

Informal schooling only

Some primary schooling

Primary schooling completed

Intermediate school or some secondary school / high school

Secondary school / high school completed

Post-secondary qualifications other than university

Some university

University completed

Post-graduate

Prefer not to say

Q4: For the following statement, please indicate whether you agree or disagree: “The government is generally responsive to demands made by people like me.” [Strongly agree, agree, neither agree nor disagree, disagree, strongly disagree, prefer not to say]

Q5: For the following statement, please indicate whether you agree or disagree: “The government should carry out its policies through specific revenue funds as opposed to involving the general revenue fund and a large state apparatus.” [Strongly agree, agree, neither agree nor disagree, disagree, strongly disagree, prefer not to say]

Q6: What is your ethnic community, cultural group, or tribe?

Akan

Ewe/Anlo

Ga/Adangbe

Dagomba

Guan

Gurma

Grusi

Mande

Other

Prefer not to say

Q7: Do you feel close to any particular political party?

Yes

No

Prefer not to say

[Answer Q7.1 if “Yes” is selected.]

Q7.1: Which party is that?

Convention People’s Party (CPP)

National Democratic Congress (NDC)

New Patriotic Party (NPP)

People’s National Convention (PNC)

Progressive People’s Party (PPP)

Democratic People's Party (DPP)
Great Consolidated Popular Party (GCPP)
National Democratic Party (NDP)
All People's Convention (APC)
Prefer not to say

CONTROL GROUP TEXT: The government of Ghana has currently two important policy priorities: 1) improving national public health in the country which would benefit people like you, and 2) increasing the collection of tax revenues from a broad set of different taxes. The government is currently considering what to do with this additional tax money.

One option the government can consider is that the money raised through the extra taxes would go to the general revenue fund managed at the discretion of the government.

TREATMENT GROUP TEXT: The government of Ghana has currently two important policy priorities: 1) improving national public health in the country which would benefit people like you, and 2) increasing the collection of tax revenues from a broad set of different taxes. The government is currently considering what to do with this additional tax money.

One option the government can consider is that the money raised through the extra taxes would go only to a dedicated government revenue fund which is responsible specifically for these public health policies, the National Health Insurance Fund (NHIF).

Q8: For each of the following activities, please indicate whether you would engage in politics [Yes, no, prefer not to say]:

Talk to others in my neighborhood

Sign a petition

Contact my local Member of Parliament

Attend a citizen meeting in my neighborhood

Join a civil society organization

Vote in upcoming elections

Join a protest

Q9: For each of the following activities, please indicate whether you would engage in politics to get the government to release information on how it will spend taxpayers' money [Yes, no, prefer not to say]:

Talk to others in my neighborhood

Sign a petition

Contact my local Member of Parliament

Attend a citizen meeting in my neighborhood

Join a civil society organization

Vote in upcoming elections

Join a protest

Q10: For each of the following activities, please indicate whether you would engage in politics to reduce the amount of freedom the government will have in the future to decide what happens with your taxes [Yes, no, prefer not to say]:

Talk to others in my neighborhood

Sign a petition

Contact my local Member of Parliament

Attend a citizen meeting in my neighborhood

Join a civil society organization

Vote in upcoming elections

Join a protest

Q11: For each of the following activities, please indicate whether you would engage in politics to change how the increased tax revenue from the government's target is spent [Yes, no, prefer not to say]:

Talk to others in my neighborhood

Sign a petition

Contact my local Member of Parliament

Attend a citizen meeting in my neighborhood

Join a civil society organization

Vote in upcoming elections

Join a protest

Q12: Reflecting on your answers, why did you respond the way you did? Please write your answer in the text box below. If you prefer not to say, please write "Prefer not

to say”.

Q13: For the following statement, please indicate how likely it is to happen on a scale from 0 to 10. 0 means extremely unlikely, and 10 means extremely likely.

“A change in the next government will reverse the current spending priorities.”

Q14: For the following statement, please indicate how likely it is to happen on a scale from 0 to 10. 0 means extremely unlikely, and 10 means extremely likely.

“Others in my neighborhood will engage in politics in the future.”

Q15: For the following statement, please indicate whether you agree or disagree: “The government is generally responsive to demands made by people like me.” [Strongly agree, agree, neither agree nor disagree, disagree, strongly disagree, prefer not to say]

Q16: For the following statement, please indicate whether you agree or disagree: “Citizens in Ghana are generally left out of the political deliberation and decision-making process.” [Strongly agree, agree, neither agree nor disagree, disagree, strongly disagree, prefer not to say]

Q17: For the following statement, please indicate whether you agree or disagree: “The government is generally competent in managing the public’s finances.” [Strongly agree, agree, neither agree nor disagree, disagree, strongly disagree, prefer not to say]

Q18: For the following statement, please indicate whether you agree or disagree: “Taxpayers must pay the taxes that they owe to the government, regardless of the quality of public services.” [Strongly agree, agree, neither agree nor disagree, disagree, strongly disagree, prefer not to say]

Q19: What is the main policy priority the government wants to pursue according to the text that you just read?

Public health

Digitalization

Agricultural development

Prefer not to say

Q20: For the following statement, please indicate whether you agree or disagree: “The taxes I normally pay are closely linked to particular benefits and services I receive.” [Strongly agree, agree, neither agree nor disagree, disagree, strongly disagree, prefer not to say]

Q21: For the following statement, please indicate whether you agree or disagree: “The government does a good job at releasing transparent information about how it spends our tax money.” [Strongly agree, agree, neither agree nor disagree, disagree, strongly disagree, prefer not to say]

Q22: For the following statement, please indicate whether you agree or disagree: “The government has a lot of freedom to spend our tax money as it pleases.” [Strongly agree, agree, neither agree nor disagree, disagree, strongly disagree, prefer not to say]

C.1.2 Design Diagnostics

Table C.1: Balance Tests for Covariates

	N	Control mean	Tax earmarking	<i>SE</i>	<i>p</i>
Age	815	28.03	-0.06	0.52	0.91
Female	872	0.45	-0.02	0.03	0.47
Education (cont.)	861	7.10	0.08	0.09	0.38
Main ethnicity	859	0.50	0.02	0.03	0.66
Partisan	811	0.28	0.04	0.03	0.19
Incumbent supporter	233	0.51	0.00	0.07	0.97
Govt. responsiveness (cont.)	870	2.73	0.09	0.09	0.29
Policy benefits (cont.)	868	4.01	-0.01	0.05	0.79

Note:

We report here a series of individual OLS regressions of the main covariate of interest on the treatment indicator. These covariates include characteristics of survey respondents. These regressions include heteroskedasticity-robust standard errors.

Table C.2: Factual Manipulation Checks

	Tax-benefit linkage	Info release	Discretion
Tax earmarking	0.026 (0.080)	0.096 (0.077)	−0.031 (0.104)
Control mean	2.674	1.995	3.070
Num. obs.	872	871	866

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Coefficients on *Tax earmarking* are OLS estimates with heteroskedasticity-robust standard errors. *Tax-benefit linkage* measures respondents' agreement with the statement that taxes and benefits are closely linked. *Info release* measures respondents' agreement with the statement that the government transparently releases information about spending priorities. *Discretion* measures respondents' agreement with the statement that the government has a lot of discretion over spending policy. *Tax earmarking* is a binary indicator for whether the respondent was exposed to the earmarking prime (treatment) or the general fund prime (control).

C.2 Appendices for Main Results

Table C.3: Effect of Tax Earmarking on Political Participation: Covariate Adjustment

	Participation	Information seeking	Discretion reduction	Policy change
Tax earmarking	−0.105 (0.152)	−0.095 (0.157)	−0.130 (0.163)	−0.227 (0.156)
Control mean	2.235	3.346	3.512	4.123
Num. obs.	621	649	659	651

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Coefficients on *Tax earmarking* are OLS estimates with heteroskedasticity-robust standard errors. *Participation* is the number of political activities respondents indicated they would engage in. *Information seeking* is the number of political activities respondents indicated they would engage in to seek policy information. *Discretion reduction* is the number of political activities respondents indicated they would engage in to reduce the government's policy discretion. *Policy change* is the number of political activities respondents indicated they would engage in to change current government policy. *Tax earmarking* is a binary indicator for whether the respondent was exposed to the earmarking prime (treatment) or the general fund prime (control). Covariates included a respondent's age, gender, education, ethnicity, partisanship, as well as responses to survey questions that asked about their perception of government responsiveness and their preference for carrying out government policy through dedicated agencies.

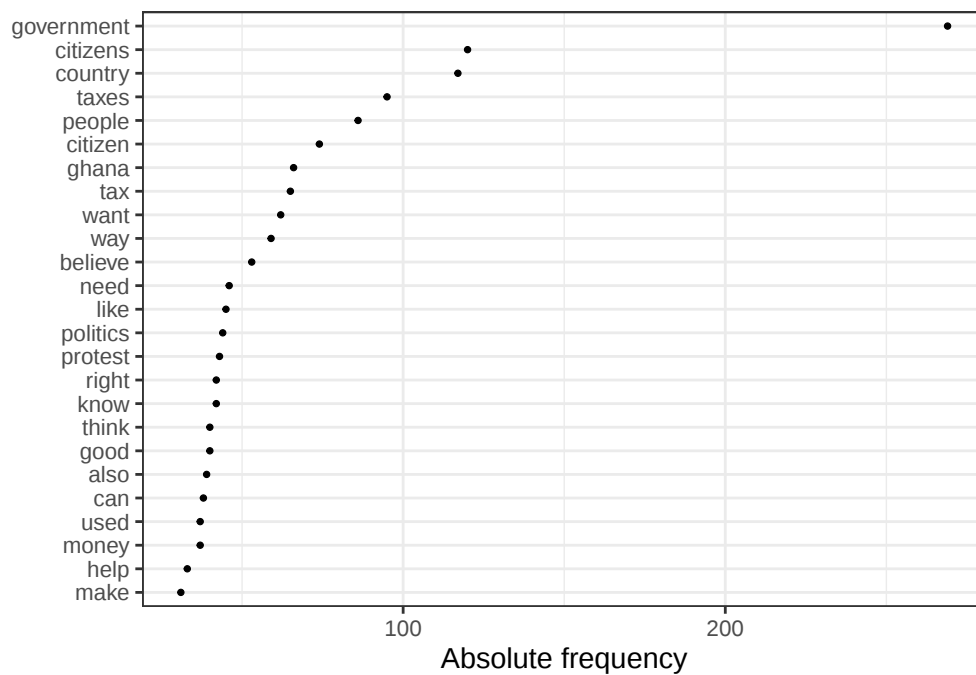
Table C.4: LDA Topic Model for Comments in Open-Ended Question

Topic 1	Topic 2	Topic 3	Topic 4	Topic 5
government	government	country	government	politics
taxes	citizens	citizen	way	like
citizens	people	want	people	political
tax	think	believe	also	vote
know	things	need	protest	power
used	feel	good	ghana	change
money	current	development	responded	make
must	needs	better	can	party
accountable	can	decision	right	government
revenue	policies	making	listen	ghana

Note:

I post-process the text data from the comments made in open-ended questions asking why participants responded the way they did into a corpus of English language comments, clean strings and remove common stopwords in English. I conduct this for all citizens who answered the question and made a comment. Based on this corpus, I perform semi-supervised Latent Dirichlet allocation (seeded-LDA) and show here the 10 most common words of 5 topics of comments.

Figure C.1: 25 Most Written Words in Open-Ended Question



Notes: I post-process the text data from the comments made in open-ended questions asking why participants responded the way they did into a corpus of English language comments, clean strings and remove common stopwords in English. I conduct this for all citizens who answered the question and made a comment. Based on this corpus, this plot shows the absolute frequency with which the top 25 most common words written in the open-ended question occurs.

C.3 Discussion of Differences with Pre-Analysis Plan

In the pre-analysis plan, power was calculated with 1,000 respondents as an input. However, due to technical difficulties in the survey implementation, only 874 respondents could be recruited. Thus, I repeat here the power calculation for a set of 874 respondents. The previous power calculation in the pre-analysis plan was as follows: “For the simple difference-in-means analysis of the effect of the treatment on the primary outcome variable Y_l , with 999 respondents we would be able to detect an effect of 0.5 points on the scale $[0, 7]$ given a standard deviation on the outcome variable of 0.282.” Here, there was a typo in the effect size: factually correct is 0.05 and not 0.5. With an adjusted number of respondents as well as an empirically calculated standard deviation of the outcome variable of 1.99, we can reformulate as follows: “For the simple difference-in-means analysis of the effect of the treatment on the primary outcome variable Y_l , with 862 respondents we would be able to detect an effect of 0.38 points on the scale $[0, 7]$ given a standard deviation on the outcome variable of 1.99.”

In the pre-analysis plan, attrition was mentioned as a contingency that would be examined across treatment groups. The pre-analysis plan read: “To assess whether attrition is systematically related to the treatments, I will measure whether the respondent completed the survey until after the primary outcomes of interest and then assess whether the probability of attrition is different across treatments.” However, this analysis could not be performed due to the fact that the survey company only delivered complete interviews conditional on giving consent to participate.

Bibliography

- Abramson, Scott F., Korhan Koçak and Asya Magazinnik. 2019. “What Do We Learn About Voter Preferences From Conjoint Experiments?”.
- Acemoglu, Daron, Camilo García-Jimeno and James Robinson. 2015. “State Capacity and Economic Development: A Network Approach.” *American Economic Review* 105:2364–2409.
- Acemoglu, Daron and James A. Robinson. 2019. *The Narrow Corridor: States, Society and the Fate of Liberty*. Penguin Press.
- Acemoglu, Daron and James A. Robinson. 2022. “Weak, Despotic, or Inclusive? How State Type Emerges from State versus Civil Society Competition.” *American Political Science Review* p. 1–14.
- Acemoglu, Daron and James Robinson. 2005. *Economic Origins of Dictatorship and Democracy*. Cambridge University Press.
- Acemoglu, Daron, James Robinson and Rafael Santos. 2013. “The Monopoly of Violence: Evidence from Colombia.” *Journal of the European Economic Association* 11:5–44.
- Acharya, Avidit, Matthew Blackwell and Maya Sen. 2016. “Explaining Causal Findings Without Bias: Detecting and Assessing Direct Effects.” *American Political Science Review* 110:512–529.
- Agerberg, Mattias. 2020. “The Lesser Evil? Corruption Voting and the Importance of Clean Alternatives.” *Comparative Political Studies* 53:253–287.
- Alatas, Vivi, Abhijit Banerjee, Rema Hanna, Benjamin A Olken and Julia Tobias. 2012. “Targeting the poor: evidence from a field experiment in Indonesia.” *American Economic Review* 102(4):1206–40.

- Anderson, Christopher J. and Pablo Beramendi. 2012. “Left Parties, Poor Voters, and Electoral Participation in Advanced Industrial Societies.” *Comparative Political Studies* 45(6):714–746.
- Anduiza, Eva, Aina Gallego and Jordi Muñoz. 2013. “Turning a Blind Eye: Experimental Evidence of Partisan Bias in Attitudes Toward Corruption.” *Comparative Political Studies* 46(12):1664–1692.
- Ansolabehere, Stephen and James M. Snyder. 2000. “Valence Politics and Equilibrium in Spatial Election Models.” *Public Choice* 103:327–336.
- Anyidoho, Nana Akua, Max Gallien, Mike Rogan and Vanessa van den Boogaard. 2022. “Mobile Money Taxation and Informal Workers: Evidence from Ghana’s E-Levy.” *ICTD Working Paper 146*.
- Aragonès, Enriqueta, Thomas Palfrey and Andrew Postlewaite. 2007. “Political Reputations and Campaign Promises.” *Journal of the European Economic Association* 5:846–884.
- Ashworth, Scott. 2012. “Electoral Accountability: Recent Theoretical and Empirical Work.” *Annual Review of Political Science* 15(1):183–201.
- Aytekin, E. Attila. 2013. “Tax Revolts During the Tanzimat Period (1839–1876) and Before the Young Turk Revolution (1904–1908): Popular Protest and State Formation in the Late Ottoman Empire.” *Journal of Policy History* 25(3):308–333.
- Bågenholm, Andreas. 2013*a*. “The Electoral Fate and Policy Impact of “Anti-Corruption Parties” in Central and Eastern Europe.” *Human Affairs* 23:174–195.
- Bågenholm, Andreas. 2013*b*. “Throwing the Rascals Out? The Electoral Effects of Corruption Allegations and Corruption Scandals in Europe 1981–2011.” *Crime, Law and Social Change* 60:595–609.
- Bågenholm, Andreas and Nicholas Charron. 2014. “Do Politics in Europe Benefit from Politicising Corruption?” *West European Politics* 37:903–931.
- Balán, Pablo, Augustin Bergeron, Gabriel Tourek and Jonathan Weigel. 2022. “Local Elites as State Capacity: How City Chiefs Use Local Information to Increase Tax Compliance in the D.R. Congo.” *American Economic Review* 112(3):762–797.

- Baldwin, Kate. 2015. *The Paradox of Traditional Chiefs in Democratic Africa*. Cambridge University Press.
- Bansak, Kirk, Jens Hainmueller, Daniel J. Hopkins and Teppei Yamamoto. 2019. "Beyond the Breaking Point? Survey Satisficing in Conjoint Experiments." *Political Science Research and Methods* pp. 1–19.
- Bardhan, Pranab. 2016. "State and Development: The Need for a Reappraisal of the Current Literature." *Journal of Economic Literature* 105:862–892.
- Basurto, Maria Pia, Pascaline Dupas and Jonathan Robinson. 2019. "Decentralization and Efficiency of Subsidy Targeting: Evidence from Chiefs in Rural Malawi." *Journal of Public Economics* p. 4047.
- Bergeron, Augustin, Elie Kabue Ngindu, Gabriel Tourek and Jonathan Weigel. 2022. "Informal Elites as Local Bureaucrats: How Working as a Tax Collector Improves the Performance of City Chiefs in the D.R. Congo." *Working Paper*.
- Besley, Timothy and Torsten Persson. 2009. "The Origins of State Capacity: Property Rights, Taxation and Policy." *The American Economic Review* 99(4):1218–1244.
- Besley, Timothy and Torsten Persson. 2010. "State Capacity, Conflict, and Development." *Econometrica* 78:1–34.
- Besley, Timothy and Torsten Persson. 2011. *Pillars of Prosperity: The Political Economics of Development Clusters*. Princeton University Press.
- Besley, Timothy and Torsten Persson. 2014. "Why Do Developing Countries Tax so Little?" *The Journal of Economic Perspectives* 28(4):99–120.
- Besley, Timothy, Torsten Persson and Daniel M. Sturm. 2010. "Political Competition, Policy and Growth: Theory and Evidence from the US." *Review of Economic Studies* 77:1329–1352.
- Blattman, Christopher, Alexandra C. Hartman and Robert A. Blair. 2014. "How to Promote Order and Property Rights under Weak Rule of Law? An Experiment in Changing Dispute Resolution Behavior through Community Education." *American Political Science Review* 108(1):100–120.
- Bodea, Cristina and Adrienne LeBas. 2016. "The Origins of Voluntary Compliance: Attitudes toward Taxation in Urban Nigeria." *British Journal of Political Science* 46(1):215–238.

- Boix, Carles and Susan C. Stokes. 2003. "Endogenous Democratization." *World Politics* 55(4):517–549.
- Bolt, Jutta, Leigh Gardner, Jennifer Kohler, Jack Paine and James A. Robinson. 2022. "African Political Institutions and the Impact of Colonialism." *Working Paper* .
- Boone, Catherine. 2003. *Political Topographies of the African State: Territorial Authority and Institutional Choice*. Cambridge University Press.
- Boone, Catherine. 2014. *Property and Political Order in Africa: Land Rights and the Structure of Politics*. Cambridge University Press.
- Breitenstein, Sofia. 2019. "Choosing the Crook: A Conjoint Experiment on Voting for Corrupt Politicians." *Research and Politics* pp. 1–8.
- Brockmann, Hilke, Philipp Genschel and Laura Seelkopf. 2016. "Happy Taxation: Increasing Tax Compliance Through Positive Rewards?" *Journal of Public Policy* 36:381–406.
- Buchanan, James M. 1963. "The Economics of Earmarked Taxes." *Journal of Political Economy* 71:457–469.
- Bueno de Mesquita, Bruce and Alastair Smith. 2011. *The Dictator's Handbook: Why Bad Behavior is Almost Always Good Politics*. Public Affairs.
- Bueno de Mesquita, Bruce, Alastair Smith, Randolph M. Siverson and James D. Morrow. 2003. *The Logic of Political Survival*. MIT Press.
- Bueno de Mesquita, Ethan and Mehdi Shadmehr. 2022. "Rebel Motivations and Repression." *American Political Science Review* p. 1–17.
- Busby, Ethan C., Joshua R. Gubler and Kirk A. Hawkins. 2019. "Framing and Blame Attribution in Populist Rhetoric." *Journal of Politics* 81:616–630.
- Castañeda, Néstor, David Doyle and Cassilde Schwartz. 2020. "Opting Out of the Social Contract: Tax Morale and Evasion." *Comparative Political Studies* 53:1175–1219.
- Centeno, Miguel. 1997. "Blood and Debt: War and Taxation in Nineteenth-Century Latin America." *American Journal of Sociology* 102:1565–1605.

- Christensen, Darin, Alexandra C. Hartman and Cyrus Samii. 2021*a*. “Citizen Monitoring Promotes Informed and Inclusive Forest Governance in Liberia.” *Proceedings of the National Academy of Sciences* 118(29):e2015169118.
- Christensen, Darin, Alexandra C. Hartman and Cyrus Samii. 2021*b*. “Legibility and External Investment: An Institutional Natural Experiment in Liberia.” *International Organization* 75(4):1087–1108.
- Cogneau, Denis, Marc Gurgand, Justine Knebelmann, Victor Pouliquen and Bassirou Sarr. 2020. “Bringing Property Owners into the Tax Net: Evidence from Dakar, Senegal.”
- de Figueiredo, Miguel F.P., F. Daniel Hidalgo and Yuri Kasahara. 2014. “When Do Voters Punish Corrupt Politicians? Experimental Evidence from Brazil.”
- de la Cuesta, Brandon, Lucy Martin, Helen Milner and Daniel Nielson. 2021. “Owning It: Accountability and Citizens’ Ownership over Oil, Aid, and Taxes.” *Journal of Politics*.
- de la Cuesta, Brandon, Lucy Martin, Helen V. Milner and Daniel L. Nielson. 2020. “Foreign Aid, Oil Revenues, and Political Accountability: Evidence from Six Experiments in Ghana and Uganda.” *Review of International Organizations* 16:521–548.
- De Vries, Catherine E and Hector Solaz. 2017. “The Electoral Consequences of Corruption.” *Annual Review of Political Science* 20:391–408.
- Degrave, Anne. 2023. “Local Rule, Elites, and Popular Grievances: Evidence from Ancien Régime France.” *Journal of Historical Political Economy* 3:1–29.
- Dhillon, Amrita and Carlo Perroni. 2001. “Tax Earmarking and Grass-Roots Accountability.” *Economics Letters* 72:99–106.
- Douglas, Mary. 1963. *The Lele of the Kasai*. Routledge.
- Duch, Raymond and Randolph Stevenson. 2008. *The Economic Vote: How Political and Economic Institutions Condition Election Results*. New York: Cambridge University Press.
- Dunning, Thad, Guy Grossman, Macartan Humphreys, Susan D. Hyde, Craig McIntosh, Gareth Nellis, Claire L. Adida, Eric Arias, Clara Bicalho, Taylor C. Boas, Mark T. Buntaine, Simon Chauchard, Anirvan Chowdhury, Jessica Gottlieb,

- F. Daniel Hidalgo, Marcus Holmlund, Ryan Jablonski, Eric Kramon, Horacio Larreguy, Malte Lierl, John Marshall, Gwyneth McClendon, Marcus A. Melo, Daniel L. Nielson, Paula M. Pickering, Melina R. Platas, Pablo Querubín, Pia Raffler and Neelanjan Sircar. 2019. “Voter Information Campaigns and Political Accountability: Cumulative Findings from a Preregistered Meta-Analysis of Coordinated Trials.” *Science Advances* 5(7):eaaw2612.
- Eggers, Andrew C. 2014. “Partisanship and Electoral Accountability: Evidence from the UK Expenses Scandal.” *Quarterly Journal of Political Science* 9(4):441–472.
- Eggers, Andrew C, Nick Vivyan and Markus Wagner. 2018. “Corruption, Accountability, and Gender: Do Female Politicians Face Higher Standards in Public Life?” *Journal of Politics* 80(1):321–326.
- Englebert, Pierre and Emmanuel Kasongo Mungongo. 2016. “Misguided and Misdiagnosed: The Failure of Decentralization Reforms in the DR Congo.” *African Studies Review* 59(1):5–32.
- Engler, Sarah. 2016. “Corruption and Electoral Support for New Political Parties in Central and Eastern Europe.” *West European Politics* 39:278–304.
- Engler, Sarah. 2020. ““Fighting Corruption” or “Fighting the Corrupt Elite”? Politicizing Corruption Within and Beyond the Populist Divide.” *Democratization* 27:643–661.
- Faguet, Jean-Paul. 2004. “Does Decentralization Increase Government Responsiveness to Local Needs? Evidence from Bolivia.” *Journal of Public Economics* 88:867–893.
- Fearon, James D. 1999. Electoral Accountability and the Control of Politicians: Selecting Good Types versus Sanctioning Poor Performance. In *Democracy, Accountability, and Representation*, ed. Adam Przeworski, Susan C. Stokes and Bernard Manin. Cambridge: Cambridge University Press pp. 55–97.
- Fernández-Vázquez, Pablo and Pablo Barberá. 2016. “Rooting Out Corruption or Rooting for Corruption? The Heterogeneous Electoral Consequences of Scandals.” *Political Science Research and Methods* 4(2):379–397.
- Fiorina, Morris P. 1981. *Retrospective Voting in American National Elections*. New Heaven: Yale University Press.

- Fisman, Ray and Miriam A. Golden. 2017. *Corruption: What Everyone Needs to Know*. New York: Oxford University Press.
- Flores-Macías, Gustavo A. 2018. “Building Support for Taxation in Developing Countries: Experimental Evidence From Mexico.” *World Development* 105:13–24.
- Franchino, Fabio and Francesco Zucchini. 2015. “Voting in a Multi-dimensional Space: A Conjoint Analysis Employing Valence and Ideology Attributes of Candidates.” *Political Science Research and Methods* 3(2):221–241.
- Fukuyama, Francis. 2011. *The Origins of Political Order: From Prehuman Times to the French Revolution*. Farrar, Straus and Giroux.
- García-Ponce, Omar, Thomas Zeitzoff and Leonard Wantchekon. 2021. “Are Voters Too Afraid to Tackle Corruption? Survey and Experimental Evidence From Mexico.” *Political Science Research and Methods* pp. 1–19.
- Garfias, Francisco and Emily A. Sellars. 2023. State Building in Historical Political Economy. In *The Oxford Handbook of Historical Political Economy*. Oxford University Press pp. 1–24.
- Gomez, Brad T. and Matthew Wilson. 2001. “Political Sophistication and Economic Voting in the American Electorate: A Theory of Heterogeneous Attribution.” *American Journal of Political Science* 45(4):899–914.
- Gottlieb, Jessica. 2022. “How Economic Informality Constrains Demand for Programmatic Policy.” *American Journal of Political Science* pp. 1–18.
- Gottlieb, Jessica and Florian M Hollenbach. 2020. “Fiscal Capacity as a Moderator of the Taxation-Accountability Hypothesis.”
- Gould, David J. 1991. Administrative Corruption: Incidence, Causes, and Remedial Strategies. In *Handbook of Comparative and Development Public Administration*, ed. Ali Farazmand. New York: Marcel Dekker pp. 467–484.
- Green, Jane. 2007. “When Voters and Parties Agree: Valence Issues and Party Competition.” *Political Studies* 55:629–655.
- Grossman, Guy, Kristin Michelitch and Carlo Prato. 2022. “The Effect of Sustained Transparency on Electoral Accountability.” *Working Paper* .

- Grossman, Guy and Tara Slough. 2022. "Government Responsiveness in Developing Countries." *Annual Review of Political Science* 25(1):131–153.
- Gruber, Jonathan. 2019. *Public Finance and Public Policy*. 6 ed. Worth Publishers.
- Hainmueller, Jens, Daniel J. Hopkins and Teppei Yamamoto. 2014. "Causal Inference in Conjoint Analysis: Understanding Multidimensional Choices Via Stated Preference Experiments." *Political Analysis* 22(1):1–30.
- Hainmueller, Jens, Dominik Hangartner and Teppei Yamamoto. 2015. "Validating Vignette and Conjoint Survey Experiments Against Real-World Behavior." *PNAS* 112(8):2395–2400.
- Hallenberg, Mats. 2009. Peasants and Tax-Farmers in Seventeenth-Century Sweden: Local Conflict and Institutional Change. In *Empowering Interactions: Political Cultures and the Emergence of the State in Europe 1300–1900*, ed. André Holenstein, Wim Blockmans and Daniel Schläppi. Routledge.
- Hanley, Seán and Allan Sikk. 2016. "Economy, Corruption or Floating Voters? Explaining the Breakthroughs of Anti-Establishment Reform Parties in Eastern Europe." *Party Politics* 22:522–533.
- Harding, Robin. 2015. "Attribution and Accountability: Voting for Roads in Ghana." *World Politics* 67(4):656–689.
- Harding, Robin. 2020. "Who Is Democracy Good For? Elections, Rural Bias, and Health and Education Outcomes in Sub-Saharan Africa." *The Journal of Politics* 82(1):241–254.
- Harding, Robin and David Stasavage. 2014. "What Democracy Does (and Doesn't Do) for Basic Services: School Fees, School Inputs, and African Elections." *Journal of Politics* 76:229–245.
- Harding, Robin and Tanushree Goyal. 2019. "Information, Attribution, and Candidate Selection: Evidence from a Survey Experiment in Ghana."
- Harris, Adam S. and Lauren Honig. 2023. "Mutual Dependence and Expectations of Cooperation." *Journal of Politics* 85(1):1–17.
- Hartman, Alexandra C., Robert A. Blair and Christopher Blattman. 2021. "Engineering Informal Institutions: Long-Run Impacts of Alternative Dispute Resolution on Violence and Property Rights in Liberia." *The Journal of Politics* 83(1):381–389.

- Haughton, Tim, Marek Rybar and Kevin Deegan-Krause. 2021. "Corruption, Campaigning, and Novelty: The 2020 Parliamentary Elections and the Evolving Patterns of Party Politics in Slovakia." *East European Politics and Societies and Cultures* pp. 1–25.
- Healy, Andrew and Neil Malhotra. 2013. "Retrospective Voting Reconsidered." *Annual Review of Political Science* 16(1):285–306.
- Henn, Soeren. 2020. "The State, Chiefs, and Development: Evidence from Africa." *Working Paper* .
- Henn, Soeren, Laura Paler, Wilson Prichard, Raul Sanchez de la Sierra and Cyrus Samii. 2022. "Engaging or Avoiding the State: How Empowerment Affects Citizen Payments for Public Services." *Working Paper* pp. 1–39.
- Hirschman, Albert O. 1972. *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States*. Harvard University Press.
- Hobolt, Sara B. and Catherine E. De Vries. 2015. "Issue Entrepreneurship and Multiparty Competition." *Comparative Political Studies* 48:1159–1185.
- Honig, Lauren. 2017. "Selecting the State or Choosing the Chief? The Political Determinants of Smallholder Land Titling." *World Development* 100:94–107.
- Honig, Lauren. 2022. *Land Politics: How Customary Institutions Shape State Building in Zambia and Senegal*. Cambridge University Press.
- Hsiung, Bingyung. 2001. "A Note on Earmarked Taxes." *Public Finance Review* 29(3):223–232.
- IMF. 2007. *Manual on Fiscal Transparency*.
- Incerti, Trevor. 2020. "Corruption Information and Vote Share: A Meta-Analysis and Lessons for Experimental Design." *American Political Science Review* pp. 1–14.
- Iversen, Vegard, Odd-Helge Fjeldstad, Godfrey Bahiigwa, Frank Ellis and Robert James. 2006. "Private Tax Collection, Remnant of the Past or a Way Forward? Evidence From Rural Uganda." *Public Administration and Development* 26(4):317–328.
- Jibao, Samuel S, Wilson Prichard and Vanessa van den Boogaard. 2017. "Informal Taxation in Post-Conflict Sierra Leone: Taxpayers' Experiences and Perceptions."

- Kallbekken, Steffen and Marianne Aasen. 2010. "The Demand for Earmarking: Results From a Focus Group Study." *Ecological Economics* 69:2183–2190.
- Kane, John V. and Jason Barabas. 2019. "No Harm in Checking: Using Factual Manipulation Checks to Assess Attentiveness in Experiments." *American Journal of Political Science* 63:234–249.
- Kasara, Kimuli. 2007. "Tax Me If You Can: Ethnic Geography, Democracy, and the Taxation of Agriculture in Africa." *American Political Science Review* 101(1):159–172.
- Key, Valdimer Orlando. 1966. *The Responsible Electorate. Rationality in Presidential Voting, 1936-1960*. Cambridge, MA: Harvard University Press.
- Khan, Adnan Q, Asim I Khwaja and Benjamin A Olken. 2016. "Tax Farming Redux: Experimental Evidence on Performance Pay for Tax Collectors." *Quarterly Journal of Economics* pp. 219–271.
- Kiser, Edgar and Margaret Levi. 2015. *Interpreting the Comparative History of Fiscal Regimes*. Cambridge University Press p. 557–571.
- Kiser, Edgar and Steven M Karceski. 2017. "Political Economy of Taxation." *Annual Review of Political Science* 20:75–92.
- Klašnja, Marko, Andrew T. Little and Joshua A. Tucker. 2016. "Political Corruption Traps." *Political Science Research and Methods* 6(3):1–16.
- Klašnja, Marko and Joshua A. Tucker. 2013. "The Economy, Corruption, and the Vote: Evidence From Experiments in Sweden and Moldova." *Electoral Studies* 32:536–543.
- Klašnja, Marko, Joshua A. Tucker and Kevin Deegan-Krause. 2014. "Pocketbook vs. Sociotropic Corruption Voting." *British Journal of Political Science* 46:67–94.
- Klašnja, Marko, Noam Lupu and Joshua A. Tucker. 2020. "When Do Voters Sanction Corrupt Politicians?" *Journal of Experimental Political Science* pp. 1–11.
- Konstantinidis, Iannis and Georgios Xezonakis. 2013. "Sources of Tolerance Towards Corrupted Politicians in Greece: The Role of Trade Offs and Individual Benefits." *Crime, Law and Social Change* 60:549–563.

- Kosack, Stephen and Archon Fung. 2014. "Does Transparency Improve Governance?" *Annual Review of Political Science* 17(1):65–87.
- Leeper, Thomas J., Sara B. Hobolt and James Tilley. 2019. "Measuring Subgroup Preferences in Conjoint Experiments." *Political Analysis* pp. 1–15.
- Leiken, Robert S. 1997. "Controlling the Global Corruption Epidemic." *Foreign Policy* (105):55–73.
- Levi, Margaret. 1988. *Of Rule and Revenue*. University of California Press.
- Levi, Margaret and Edgar Kiser. 2015. "Spend and Tax: Why and When States Underprovide Public Goods." *Working Paper* .
- Lieberman, Evan and Yang-Yang Zhou. 2021. "Self-Efficacy and Citizen Engagement in Development: Experimental Evidence from Tanzania." *Journal of Experimental Political Science* pp. 1–18.
- Lockerbie, Brad. 2008. *Do Voters Look to the Future? Economics and Elections*. New York: SUNY Press.
- Lust, Ellen and Lise Rakner. 2018. "The Other Side of Taxation: Extraction and Social Institutions in the Developing World." *Annual Review of Political Science* 21:277–294.
- Mamdani, Mahmood. 1996. *Citizen and Subject*. Princeton University Press.
- Mares, Isabela and Giancarlo Visconti. 2020. "Voting for the Lesser Evil: Evidence From a Conjoint Experiment in Romania." *Political Science Research and Methods* 8:315–328.
- Martin, Isaac William, Ajay K. Mehrotra and Monica Prasad, eds. 2009. *The New Fiscal Sociology*. Cambridge University Press.
- Martin, Isaac William and Heather Harper. 2021. "What Makes a Tax Policy Popular? Predicting Referendum Votes from Policy Text." *Socius* 7:1–17.
- Martin, Lucy. 2016. "Taxation, Loss Aversion, and Accountability: Theory and Experimental Evidence for Taxation's Effect on Citizen Behavior." .
- McCleary, William. 1991. "The Earmarking of Government Revenue: A Review of Some World Bank Experience." *The World Bank Research Observer* 6:81–104.

- Meadowcroft, John. 2011. "Accountable and Responsive Government: The Value of Earmarked Taxation." *Economic Affairs* .
- Migdal, Joel S. 1988. *Strong Societies and Weak States: State-Society Relations and State Capabilities in the Third World*. Princeton University Press.
- Mookherjee, Dilip. 2006. "Decentralization, Hierarchies, and Incentives: A Mechanism Design Perspective." *Journal of Economic Literature* 44(2):367–390.
- Mosley, Paul and Abdul-Gafaru Abdulai. 2020. "The Political Economy of Progressive Fiscal Contracts in Africa and Latin America." *Development Policy Review* 38:411–427.
- Muñoz, Jordi, Eva Anduiza and Aina Gallego. 2016. "Why do Voters Forgive Corrupt Mayors? Implicit Exchange, Credibility of Information and Clean Alternatives." *Local Government Studies* 42(4):598–615.
- Nguema, Rano-Michel. 2005. "Développement de la Ville, Découpage et Appropriation des Territoires Urbains au Gabon: Le Cas de Libreville." *Revue Belge de Géographie* (4):481–498.
- North, Douglass C. 1991. "Institutions." *Journal of Economic Perspectives* 5(1):97–112.
- North, Douglass C and Barry R Weingast. 1989. "Constitutions and Commitment: The Evolution of Institutions Governing Public Choice in Seventeenth-Century." *Journal of Economic History* 49:803–832.
- Nzongola-Ntalaja, Georges. 1975. "Urban Administration in Zaire: A Study of Kananga, 1971–1973." *Unpublished Ph. D. diss., University of Wisconsin-Madison* 263.
- OECD. 2020. *Government at a Glance: Latin America and the Caribbean 2020*. OECD Publishing.
- Olken, Benjamin and Rohini Pande. 2019. "Governance Initiative Review Paper." *J-PAL Governance Initiative* pp. 1–118.
- Olson, Mancur. 1971. *The Logic of Collective Action: Public Goods and the Theory of Groups*. Harvard University Press.

- Olson, Mancur. 1993. "Dictatorship, Democracy, and Development." *American Political Science Review* 87(3):567–576.
- Paler, Laura. 2013. "Keeping the Public Purse: An Experiment in Windfalls, Taxes, and the Incentives to Restrain Government." *American Political Science Review* 107:706–725.
- Paler, Laura, Leslie Marshall and Sami Atallah. 2018. "The Social Costs of Public Political Participation: Evidence from a Petition Experiment in Lebanon." *Journal of Politics* 80(4):1–6.
- Paler, Laura, Wilson Prichard, Raul Sanchez de la Sierra and Cyrus Samii. 2017. "Survey on Total Tax Burden in the DRC, Final Report." *UK DFID* pp. 1–127.
- Pavão, Nara. 2018. "Corruption as the Only Option: The Limits to Electoral Accountability." *The Journal of Politics* 80(3):996–1010.
- Pomeranz, Dina and José Vila-Belda. 2019. "Taking State-Capacity Research to the Field: Insights from Collaborations with Tax Authorities." *Annual Review of Economics* 11.
- Pontusson, Jonas and David Rueda. 2010. "The Politics of Inequality: Voter Mobilization and Left Parties in Advanced Industrial States." *Comparative Political Studies* 43(6):675–705.
- Prazniak, Roxann. 1980. "Tax Protest at Laiyang, Shandong, 1910: Commoner Organization Versus the County Political Elite." *Modern China* 6(1):41–71.
- Prichard, Wilson. 2009. "The Politics of Taxation and Implications for Accountability in Ghana 1981–2008." *IDS Working Paper 330* .
- Prichard, Wilson. 2010. "Taxation and State Building: Towards a Governance Focused Tax Reform Agenda." *IDS Working Paper 341* .
- Prichard, Wilson. 2015. *Taxation, Responsiveness and Accountability in Sub-Saharan Africa*. Cambridge University Press.
- Ross, Michael L. 2001. "Does Oil Hinder Democracy?" *World Politics* 53(3):325–361.
- Sanchez de la Sierra, Raul. 2020. "On the Origins of the State: Stationary Bandits and Taxation in Eastern Congo." *Journal of Political Economy* 128:32–74.

- Scheve, Kenneth and David Stasavage. 2012. "Democracy, War, and Wealth: Lessons from Two Centuries of Inheritance Taxation." *American Political Science Review* 106:81–102.
- Schumpeter, Joseph A. 1918. "The Crisis of the Tax State."
- Scott, James C. 1990. *Domination and the Arts of Resistance: Hidden Transcripts*. Yale University Press.
- Sheely, Ryan. 2015. "Mobilization, Participatory Planning Institutions, and Elite Capture: Evidence from a Field Experiment in Rural Kenya." *World Development* 67:251–266.
- Sniderman, Paul M. 2018. "Some Advances in the Design of Survey Experiments." *Annual Review of Political Science* 21:259–277.
- Solaz, Hector, Catherine E. De Vries and Roosmarijn de Geus. 2019. "In-Group Loyalty and the Punishment of Corruption." *Comparative Political Studies* 52:896–926.
- Stasavage, David. 2020. *The Decline and Rise of Democracy: A Global History from Antiquity to Today*. Princeton University Press.
- Stephenson, Matthew C. 2015. Corruption and Democratic Institutions: A Review and Synthesis. In *Greed, Corruption, and the Modern State*, ed. Susan Rose-Ackerman and Paul Lagunes. Cheltenham: Edward Elgar pp. 92–133.
- Stokes, Donald E. 1963. "Spatial Models of Party Competition." *American Political Science Review* 57(2):368–377.
- Stokes, Susan C, Thad Dunning, Marcelo Nazareno and Valeria Brusco. 2013. *Brokers, Voters, and Clientelism: The Puzzle of Distributive Politics*. Cambridge University Press.
- Sung, Hung-En and Doris Chu. 2003. "Does Participation in the Global Economy Reduce Political Corruption?" *International Journal of Comparative Criminology* 3:94–118.
- Teja, Ranjit S. 1988. "The Case for Earmarked Taxes." *Staff Papers (International Monetary Fund)* 35:523–533.

- Tilly, Charles. 1992. *Coercion, Capital, and European States, AD 990-1992*. Blackwell Oxford.
- Timmons, Jeffrey F. 2005. "The Fiscal Contract: States, Taxes, and Public Services." *World Politics* 57:530–567.
- Tsai, Lily L. 2007. "Solidary Groups, Informal Accountability, and Local Public Goods Provision in Rural China." *American Political Science Review* 101(2):355–372.
- Tucker, Joshua A. 2006. *Regional Economic Voting: Russia, Poland, Hungary, Slovakia and the Czech Republic, 1990–1999*. Cambridge: Cambridge University Press.
- van den Boogaard, Vanessa. 2021. *Informal Revenue Generation and the State: Evidence from Sierra Leone*. Mimeo.
- van den Boogaard, Vanessa, Wilson Prichard, Rachel Beach and Fariya Mohiuddin. 2022. "Enabling Tax Bargaining: Supporting More Meaningful Tax Transparency and Taxpayer Engagement in Ghana and Sierra Leone." *Development Policy Review* 40(1):e12543.
- Vansina, Jan. 1990. *Paths in the Rainforests: Toward a History of Political Tradition in Equatorial Africa*. The University of Wisconsin Press.
- Vansina, Jan. 2010. *Being Colonized: The Kuba Experience in Rural Congo, 1880–1960*. The University of Wisconsin Press.
- Wagner, Richard. 1991. *Charging for Government: User Charges and Earmarked Taxes in Principle and Practice*. Routledge.
- Wang, Yuhua. 2021. "State and Society 2.0: Toward Fourth-Generation Theories of the State." *Comparative Politics* 54:175–198.
- Webber, Carolyn and Aaron Wildavsky. 1986. *A History of Taxation and Expenditure in the Western World*. Simon and Schuster.
- Weigel, Jonathan and Elie Kabue Ngindu. 2021. "The Taxman Cometh: Pathways out of a Low-Capacity Trap in the D.R. Congo."
- Weigel, Jonathan L. 2020. "The Participation Dividend of Taxation: How Citizens in Congo Engage More with the State when it Tries to Tax Them." *The Quarterly Journal of Economics* 135(4):1849–1903.

- Wilfahrt, Martha. 2022. *Precolonial Legacies in Postcolonial Societies: Representation and Redistribution in Decentralized West Africa*. Cambridge University Press.
- Wilkinson, Margaret. 1994. "Paying for Public Spending: Is There a Role for Ear-marked Taxes?" *Fiscal Studies* 15:119–154.
- Winters, Matthew S and Rebecca Weitz-Shapiro. 2013. "Lacking Information or Con-doning Corruption When Do Voters Support Corrupt Politicians?" *Comparative Politics* pp. 418–436.
- Xezonakis, Georgios, Spyros Kosmidis and Stefan Dahlberg. 2015. Can Elections Combat Corruption? Accountability and Partisanship. In *Elites, Institutions and the Quality of Government*, ed. Carl Dahlström and Lena Wängnerud. Basingstoke: Palgrave Macmillan UK pp. 283–304.
- Xezonakis, Georgios, Spyros Kosmidis and Stefan Dahlberg. 2016. "Can Electors Combat Corruption? Institutional Arrangements and Citizen Behaviour." *Euro-pean Journal of Political Research* 55:160–176.
- Yair, Omer, Raanan Sulitzeanu-Kenan and Yoav Dotan. 2020. "Can Institutions Make Voters Care about Corruption?" *The Journal of Politics* 82:1–13.